NASA TECHNICAL TRANSLATION

NASA TT F-119

NASA TT F-119

PROBLEMS OF COSMOGONY

VOL. VIII

by V. A. Ambartsumyan, et al.

Akademiya Nauk SSSR, Astronomicheskiy Sovet,
Izdatel'stvo Akademii Nauk SSSR, Moscow, 1962

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • MAY 1965

PROBLEMS OF COSMOGONY

VOL. VIII

By V. A. Ambartsumyan, et al.

Translation of "Voprosy kosmogonii, Tom VIII."
Akademiya Nauk SSSR, Astronomicheskiy Sovet,
Izdatel'stvo Akademii Nauk SSSR, Moscow, 1962.

Electro-Optical Systems, Inc. translated and prepared the reproducible copy for this publication and is responsible for the accuracy and fidelity of the translation and editorial quality of the contents.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – Price $6.00

TABLE OF CONTENTS

SELECTIONS OF PAPERS DELIVERED AT CONFERENCE
ON EXTRAGALACTIC ASTRONOMY
Moscow, June 1961

PROBLEMS OF EXTRAGALACTIC RESEARCH /3*

By

V. A. Ambartsumyan

Basic results of recent studies of extragalactic objects are reported. The phenomenon of superposition of subsystems inside galaxies has been considered in detail in connection with a cosmogonical interpretation.

Attention is paid to the role of galactic nuclei as the centers of cosmogonical activity. Extremely high cosmogonical activity of central nuclei in super giant galaxies is emphasized. One of the phenomena which resulted from this activity is the radio galaxies. The superposition of subsystems in clusters of galaxies is also discussed. A number of problems with which astronomers are faced in the field of extragalactic research are enumerated.

In this report, basic facts of extragalactic astronomy are discussed. Since the valid concept of outer stellar systems, or galaxies, was established scientifically only about 40 years ago, many fundamental problems concerning the universe outside our galaxy remain unsolved. Therefore, only the problems which seem to be essential for further extragalactic research are presented. We shall try not to stray too far from the facts while touching briefly on those problems whose solution seem feasible in the future with the help of available means.

As is known, extragalactic astronomy borders on cosmology which attempts to describe the universe as a whole by means of theories. These theories have undoubtedly been beneficial since certain solutions to the equations of the general theory of Einsteinian gravity have /4 been investigated using these theories, and the problem has been posed of comparing these solutions with the characteristics of an observed part of the universe. At the same time, the solutions have often served as the arena for very rough simplifications and hasty extrapolations.

In this report we will not consider the analysis of these theories or the problem of their further development, although we realize that a critical review of the work being accomplished in this region would be

---

*Numbers given in the margin indicate the pagination in the original foreign text.

extremely valuable. Nevertheless, the facts and problems which are approached below should be of value also for cosmological theories.

I. Principal Facts Concerning the Distribution of Matter

One of the properties of the universe that surrounds us is the fact that the greatest portion of matter that we have observed is centered in the stars. Other objects contain only a small part of the total observed mass.

The most important fact of extragalactic astronomy is that an overwhelming majority of the observed stars enters into the composition of giant stellar systems which bear the name galaxies.

The dimensions of the galaxies and the number of their star population are extremely varied. The super giant galaxies of the type of two of the brightest galaxies, which are found in the center of the cluster of Coma Berenices (NGC 4874 and NGC 4889), have photographic absolute magnitudes approaching -22^m and contain hundreds of billions of stars, while the dwarf systems of the type of galaxy found in Sculptor have absolute magnitudes on the order of $-11^m.0$ and evidently contain only a few million stars. However, systems of a still lower illumination, which can be called sub-dwarf galaxies, border on the dwarf galaxies. A good representative of such systems is the galaxy Capricorn, discovered by Zwicky, and which has an absolute photographic magnitude of $-6^m.5$. It should be assumed that this system contains several tens of thousands of stars at the most. Thus, this system is more than 10 million times smaller than the super giant galaxies; furthermore, as far as the number of stars is concerned, it remains far behind many globular stellar clusters.

As a rule, the diameters of the galaxies lie between the limits from 50,000 parsecs for the super giants and up to 500 parsecs for the sub-dwarfs.

Giant and super giant galaxies with diameters of 5,000 to 50,000 parsecs invariably have a high surface brightness (more than /5

$24^m.0$ per square second of arc) and they also have a great concentration of luminosity toward the center. Among dwarf galaxies, objects having high surface brightness are found along with objects of low surface brightness. Essentially however, there are among the dwarf galaxies

systems whose gradient is very small as well as systems having a large surface brightness gradient from their borders to the center. On photographs, such a system appears to be an almost uniform disc.*

The fact that the overwhelming majority of stars is included in the contents of the galaxies is of great significance, if we take into consideration the fact that galactic systems appear to be isolated from one another in the first approximation. Normally, the distance between neighboring galaxies exceeds the diameters of their central, most dense parts by many times. In addition, parts of the galaxies that are remote from the center and extremely rarefied sometimes interpenetrate each other. Along with this topographic isolation, a dynamic seclusion of the galaxies as well as that of stellar systems should be noted. By dynamic seclusion is meant that property by which the motion of the stars in each galaxy is basically determined by the total combined interaction with other members of the same galaxy. In addition, it should be noted that this condition of dynamic seclusion is accomplished only in a certain approximation. The mutual disturbances of stellar systems which are close to one another and the ejections from central parts of galaxies which will be discussed later are incidences where the dynamic seclusion of the systems and galaxies appears to be violated, more or less.

---

*The dwarf star systems of Sculptor and Fornax, discovered by Shapley, proved to be examples of galaxies which possess a small density gradient and enter into the Local Group. The surface brightnesses of these systems are unusually low. Later on, Baade showed that the galaxies NGC 147 and NGC 185, which belong to the Local Group, also have a small density gradient. The surface brightness of this pair of galaxies is significantly higher than that of the systems of Sculptor and Fornax. The other two members of the Local Group have an intermediate magnitude of surface brightness: B Sextantis ($9^h\ 57^m.3 + 5°34'$,1950) and Leonis ($10^h\ 05^m.8 + 12°33'$, 1950). In addition, their density gradient is also very small. In the Virgo cluster, there are a large number of objects of low surface brightness and small density gradient. As far as linear dimensions are concerned, some of these approach the average galaxies in size. For example, the galaxy IC 3475 in the cluster of Virgo possesses an insignificant density gradient as well as an extremely low surface brightness, and its diameter approaches 5,000 parsecs. Thus, this galaxy, as far as its dimensions are concerned, far exceeds analogous objects of the Local Group. Nevertheless, it is emphasized that objects which are relatively large and which have small density gradients and low surface brightness are very rare. For example, in the well known cluster of Cancer, the largest similar galaxy has a linear diameter of about 2500 parsecs.

Stars make up the composition of the galaxies and, in similar /6 fashion, galaxies form the composition of such systems of galaxies as the clusters of galaxies, groups of galaxies, and multiple galaxies.

If it had been assumed two decades ago that besides clusters and groups of galaxies, there exists also a general field which includes a majority of galaxies (similar to the general stellar field, which is present in our stellar system, and which includes all clusters and associations), then at the present time, the very existence of a general field would be in great doubt.

The clusters which we have observed are divided into two types: spherical clusters with a regular symmetrical distribution of the galaxies near the center and clusters of an irregular form. The population of spherical clusters consists basically of elliptical galaxies. The distribution of this cluster contains a high percentage of spirals. To the scattered clusters groups of galaxies which are similar to the Local Group or Groups around M 101 and M 81 adhere closely.

Thus, for example, the groups of galaxies associated with M 101 amd M 81 in effect do not contain an elliptical galaxy. They are composed only of spirals and irregular galaxies. The group of galaxies in Sculptor, which was investigated by de Vaucouleurs, contains only irregular gaalxies and those of the type Sc. Our Local Group also contains no elliptical galaxies of high luminosity, but it does contain some elliptical galaxies of low and moderate luminosity.

It is also interesting to note that our Local Group is essentially composed of two very small groups approaching multiple galaxies in size. The first group contains our galaxy, two Magellanic clouds, and apparently certain galaxies of the type of system in Sculptor. The second group contains the Andromeda nebula with its four satellites and M 33. However, such a division can be established only for the galaxies of a high and moderate luminosity. The possibility is not excluded that all of the space of the Local Group is occupied by dwarf galaxies. We should add that the entire mass of the Local Group is determined mainly by two galaxies which are essentially the center of these two sub-groups; that is, with a mass M 31 and the mass of our galaxy. In turn, the rich clusters of galaxies containing a large number of members sometimes are found in two's and three's, forming multiple clusters of galaxies. It was previous indicated that /7 galaxies usually proved to be isolated from one another by stellar systems. However, the case when this isolation is infringed merits our attention. Let us mention three categories of similar objects:

(a) Interaction of the galaxy. Those cases when two galaxies are found near one another, and the presence of one greatly influences the structure of the other galaxy. Numerous examples of interacting galaxies are given in the Vorontsov-Vel'yaminov Atlas. From this, two interpretations of the observed interactions are possible: (1) the tidal effect and (2) the division of two galaxies which developed together. The "interaction" observed in the last case should be considered as results of the process of division.

(b) Pairs of galaxies connected by bridges or by crosspieces. Numerous examples of this type are given in the articles of Zwicky. Experiments conducted by the latter showed that the indicated crosspieces are composed of stars.

Together with the crosspieces, jets are observed leaving the central regions of certain spherical galaxies which contain, in themselves, a bluish condensation appearing to be dwarf galaxies.* It is established that the jet somehow connects a large galaxy with dwarf galaxies resembling a crosspiece. In these cases it should not be doubted that the dwarf galaxy separated from the central nucleus of the main galaxy. Therefore, it seems more likely that the bridges and crosspieces are, in general, the result of the genetic process of two galaxies arising from one.

(c) Radio galaxies. The assumption has been expressed that radio galaxies are the result of a random collision of a pair of independent stellar systems. It was assumed that the energy of radio emission has as its source the energy from the collision of two gaseous masses respectively entering into each of the galaxies. However, the facts contradict this hypothesis. All the data point toward the fact that radio galaxies are certain very short stages in the process of interdevelopment of galaxies of very high luminosity (super giant galaxies).

Apparently the radio activity of the galaxies is closely related to new formations of the type of condensations and jets (ejected /8
from the center), spiral arms, and even of entire new galaxies. In other words, in certain cases, the process of division of the nucleus of the galaxy and the development of a new galaxy within the old occurs.

---

*NGC 3561 and IC 1182 are attributed to a number of galaxies of high luminosity from whose central regions the jets originate containing, in itself, the bluish condensation.

Therefore, radio galaxies are often over-crowded systems composed of the old galaxies and the new formations. These new formations are usually observed still submerged in the old galaxy. It should be mentioned that the above ennumerated disruptions to the isolation of the galaxies is observed only in the case of a small part of the total number of galaxies. It is likely that these disruptions only occur in the determined stage of the development of galaxies when new galaxies arise.

In spite of the fact that real progress was accomplished in the study of spatial distribution of galaxies, many basic questions remain unanswered. We will mention some of them:

a) Do the clusters of galaxies, in turn, form systems of a higher order of the type of super clusters or super galaxies?

It is certain that our Local Group is a component part of a certain group of clusters in the center of which, as a part of its nucleus, is the large aggregation in Virgo. This large spatial grouping was called a super galaxy by de Vaucouleurs. Its dimensions are on the order of 20 million parsecs. In the meantime, however, we can say nothing about the dynamic unity of this system or of the presence of forces which might support this unity.

In addition, it is very interesting to note that there is no evidence of the existence of a large number of similar super galaxies from the study of the distribution of galaxies on the celestial sphere. In the examination of this problem it is necessary to consider the existence of two possibilities: 1) the intervals between super galaxies are great as compared with the diameters of the super galaxies themselves. 2) These distances are of the same order as the diameters of the super galaxies.

In the first case, many such super galaxies must be precisely observed as isolated formations in their projection on the celestial sphere. In the second case, only a small number of similar formations will be observed from their projection as isolated systems. Since we can only scratch the surface of the problem, it will be difficult to draw a conclusion concerning the existence of distant super galaxies.

The observations give a direct indication of the irregularity in the distribution of the clusters and groups of galaxies. To a certain degree these can be explained by the existence of super galaxies. In addition, it is possible to consider that we observe not far from

us only a few isolated clouds consisting of a great number of condensations. Thereby it was established reliably that only the existence of the large cloud in the southern sky ranges from $\ell = 160^\circ$ to 240° when $b = -40^\circ$.*

From these two facts we assume the second alternative; i.e., super galaxies exist but the distance between them is approximately the /9
same as their diameters.

---

*The irregularity in the distribution of galaxies throughout the sky, apart from that which is caused by the absorption in our galaxy, is clearly shown in the case of galaxies of the catalogue of Shapley and Ames (specific magnitude $13^m.0$). This irregularity is primarily connected with the existence of a local super galaxy. The irregularity is even more clearly expressed in the results of the calculations of Shane and Wirtanen (specific magnitude $18^m.4$). Thereby, the non-homogeneities are stipulated on a small scale by a concentration of galaxies in clusters. However, there are larger homogeneities which are caused by a tendency of the clusters to form groups which are similar to the super galaxies previously discussed.

In accordance with the data of Zwicky and other authors, irregularities in the distribution of galaxies are extended to the limits that are attainable with the help of the Palomar Schmidt type telescope (almost to 20^m). It is possible, as an example, to point out large clouds of galaxies in the region of the cluster of Corona Borealis. However, for the study of the tendency of clusters to congregate, the research on the distribution of the centers of galactic clusters is of great interest. Such research was conducted by Able according to photographs of the Palomar Atlas. His results confirmed the non-homogeneity in the distribution of clusters.

Zwicky considers the main cause of the non-homogeneity observed in the distribution of clusters, i.e., the flocculent structure of the absorbent intergalactic dust. His arguments in favor of its presence in specific directions of intergalactic absorption evidently were convincing. However, not all of the deviation from homogeneity can be explained by this. Therefore, it is necessary to consider the true irregularity in the distribution of galaxies farthest away from us.

Although the existence of individual galaxies must be accepted, the following questions remain unclarified: What percentage of the clusters of galaxies enters into these systems of higher order? Is the tendency of two known types of clusters (spherical and diffuse) to congregate identically strong? It will be possible to answer these questions only on the basis of more detailed photometric and statistical investigations.

b) To what degree do galaxies of low luminosity duplicate the spatial distribution of galaxies of high luminosity? As was previously indicated, the concentration of galaxies in clusters is well established for objects of high luminosity. However, objects of low luminosity, beginning with distances of several million parsecs, will definitely be lost among galaxies in the remote background, and the solution of /10
the problem relative to them encounters known difficulties. However, concerning one class of objects of low luminosity, primarily galaxies of a low surface brightness, there are some things that can be concluded on the basis of the results of the work of Reyves who established that the distribution of objects of low surface brightness in the cluster Virgo roughly duplicates the distribution of galaxies of high luminosity. On the other hand, we cannot say whether the galaxies of low surface brightness (systems of the type in Sculptor or Zwicky's object in Capricorn) contain a general Metagalactic field or are concentrated in clusters and groups.

c) The super galaxies which were previously mentioned, are objects with diameters of approximately 20,000,000 parsecs. If they appear to be the largest inhomogeneities in the distribution of galaxies, it might be expected that spatial cells with dimensions of 50 or 100,000,000 parsecs will be approximately equal in the quantity of matter (galaxies) contained in them. However, it is possible that inhomogeneities exist on a much larger scale. This problem can be resolved only on the basis of research on the distribution of weak clusters of galaxies (up to a magnitude of 21^m) or on the basis of investigation on the distribution of extragalactic sources of radio sources. The solution of this problem is extremely important for building the foundation for those other cosmological theories. At the present time it is only possible to state that there are no indications that the cosmologist's postulate concerning the homogeneity is justified.

d) It was previously mentioned that there is good evidence in favor of the existence of intergalactic dust matter. In connection with this, it follows that there it is desirable to investigate all aspects of intergalactic matter. It is now possible to discuss the reality of certain of these forms:

1. The bright intergalactic matter sometimes filling the central portion of the space occupied by galactic

clusters. All data indicate that this bright substance, as well as the bridges and crosspieces often observed in pairs of galaxies, is composed of stars.

2. Intergalactic globular clusters. Certain of these were observed at distances greater than 100,000 parsecs.

3. Giant clouds of relativistic electrons ejected from the interiors of galaxies. For example, the radio source α-Centauri consists of three similar clouds and α-Cygnus consists of two. The dimensions of each of these clouds exceeds the dimensions of normal galaxies. Many of these clouds undoubtedly were dispersed in intergalactic space.

4. Absorbent dust matter. There are no data on the dimensions of individual clouds of dust masses.

5. The neutral gaseous masses, which are however, present in such small quantities that the radiation which they emit (for example, in the line λ = 21 cm) has not been successfully observed. /11

There is no doubt that each of these forms of intergalactic matter merits special investigation.

II. Principal Facts Concerning Kinematics and Dynamics in the Systems of Galaxies

Our knowledge of motions in the universe of galaxies is limited to information on the radial velocities of approximately 1,000 galaxies. We have no information at all concerning tangential velocities. However, existing data on radial velocities which have been obtained almost entirely by the Mount Wilson, Palomar, and Lick Observatories, confronts us with the most difficult problems ever encountered in astronomy.

The sum total of observed galaxies is a part of a more grandiose system which we call the Metagalaxy. This concept of a Metagalaxy is valid independently of the solution to the problem concerning the existence of galaxies outside this system. The expansion of the Metagalaxy is the most important factor established on the basis of our information concerning radial velocities of galaxies.

Hubble's law, derived from empirical data

$$v_r = H \cdot r,$$

is maintained with accuracies to small deviations for the values of r approaching almost 2 billion parsecs. This law relates the approximate homogeneity of the observed expansion. All attempts to find some other explanation (other than the Doppler effect) for the red shift proved to be artificial and unsuccessful. Therefore, in the examination of any problem concerning the nature, and especially the evolution, of the Metagalaxy the expansion phenomenon must be taken into consideration.

Of course, Hubble's law is only valid on the average. In addition to the velocity determined according to Hubble's formula, each cluster of galaxies and each galaxy in relation to its center of gravity has its own characteristic velocities.

Thus, in the Local Group where distances between galaxies are small, relative velocities are determined mainly by the characteristic motion of the individual members. However, the closest clusters of galaxies and the closest outer groups recede from us. This testifies to the small magnitude of characteristic velocities of these clusters and groups as compared to the systematic velocities of their recession according to Hubble's formula.

The numerical value of the constant, H, is quite significant /12 since the knowledge of it permits the determination of the distance to the most remote clusters. Unfortunately, this value is not accurately known. According to the results of Sandage (1958) the probability is high that it lies somewhere within the limits of

$$60 \text{ km/sec} \cdot \text{M parsecs} < H < 140 \text{ km/sec} \cdot \text{M parsecs}$$

and with a certain risk that it is included between

$$70 \text{ km/sec} \cdot \text{M parsecs} < H < 100 \text{ km/sec} \cdot \text{M parsecs}$$

The problems related to the determination of H are not considered. It is merely stated that, under all conditions, Hubble's law permits a good evaluation of the relative distances.

The second important fact concerning the motion of galaxies is the presence of some dispersion in the velocities of the galactic clusters which is connected to the internal motions in these clusters.

If a cluster is found in a stationary state or, after the lapse of a certain time, should enter a stationary state, then its total energy, E, should be negative:

$$E = T + U < 0,$$

where T and U are, respectively, the kinetic and the potential energy of the system. If $E > 0$ then the system cannot enter into a stationary state and at least part of its members must approach infinity.

Recent investigations have shown that, for certain groups and multiple systems, the kinetic energy of internal motion, which is determined according to the radial velocities exceeds many times the probable value of the absolute magnitude of potential energy calculated assuming that the fundamental mass of the cluster is concentrated in its galaxies and that the relation of the mass to the luminosity is $f = M/L$ for a given type of galaxy of the same order as in those cases where this relationship is determined on the basis of investigations of the rotation of galaxies. From this came the conclusion that certain groups and clusters have positive energy and should disperse in space. It was necessary to make such a deduction, for example, relative to the clusters of the galaxies in Virgo and Hercules, and for a comparatively near group in Sculptor. The latter case, analyzed in detail by de Vaucouleurs, was especially striking since the kinetic energy exceeded the estimated absolute value of the potential energy by a factor of one and a half or two. Since positive energy should lead to the /13 recession of a part of the members of the cluster, and sometimes the complete dispersion of the cluster, it is plausible to imagine that there is something in common between the phenomenon of non-stationary clusters on one hand and the phenomenon of the expansion of Metagalaxies on the other.

The systems of the type of the Local Supergalaxy should play an intermediate role in this attitude. As is known, its component parts recede from one another. For example, the cluster in Virgo, or a group connected with M 81 is receding from the Local Group of galaxies.

What was said concerning the total inner energy of the clusters of galaxies also remains correct relative to the multiple systems. Evidently, some multiple systems have over-all positive energies. These systems are from 10^9 to $5 \cdot 10^9$ years old.

However, regardless of the sign of the total energy, one more peculiarity of all multiple (triple, quandruple, etc.) galaxies draws attention to itself. As is known, an overwhelming majority of multiple stars has configurations of the "ordinary" type, while the configurations of the "Trapezium of Orion" type comprise an insignificant percentage (about 10%). Among the multiple galaxies approximately half of the systems have configurations of the Trapezium type. Since the systems of the Trapezium type as a rule are unstable, we can

conclude that the time which has passed since the formation of these multiple groups does not exceed by more than several times the period of its conversion into such a multiple system, which in turn is about 10^9 to $5 \cdot 10^9$ years old.

Finally, it should be mentioned that assumption of the negativity of the energy of all double galaxies sometimes leads to incredibly large values of the component masses (Page). Therefore, there is a basis for assuming that certain double galaxies also have positive energy.

In the overcrowded systems, such as the radio galaxies, significant diversities of velocities of the components are observed. For example, in the radio galaxy Perseus A, this difference reaches 3,000 km/sec, and these pairs therefore possess positive energy. In our opinion, we consider the formation of such a pair as from a single galaxy.

The further accumulation of data on the radial velocities of galaxies permits the solution of many unsolved problems concerning their kinematics and dynamics. Certain of these unsolved problems are described below.

a) A more accurate determination of the constant in the law of the red shift. This designates an increase in accuracy for the scale of extragalactic distances.

b) The determination of the nature of the dependence of the /14 red shift on distance with very large values for the latter. Undoubtedly we must observe the violation of linear dependency. However, for the solution of fundamental cosmological problems, it is extremely important to know on which side the tendency to stray from linearity occurs, and whether the magnitude of this tendency is dependent of direction.

c) It is very important to determine the characteristic velocities of the centers of gravity of individual clusters of galaxies; that is, their tendency to stray from the observed velocities taken from the Hubble formula. This has essential significance for the solution of problems concerning genetic bonds between neighboring clusters. As for the determination of the indicated tendencies to stray, it is necessary to be able to determine more accurately the distances of remote clusters independent of Hubble's law.

d) For the solution of many problems on the dynamics of the clusters of galaxies and multiple galaxies it is necessary to be able to determine the masses of the latter. Unfortunately, in the case of

remote galaxies entering the indicated systems, we determine the masses statistically surmising the negativity of the energy, and also the applicability of the virial theorem. It is necessary to determine the masses of the galaxies within the closest clusters independently from this supposition. In addition, it is necessary to find the extent to which the evaluations are valid, at least insofar as the maximum magnitude is concerned, in the possible intergalactic masses of each system (cluster or group).

e) The disparity between the mass of a system determined from the virial theorem and the mass found from the evaluation of individual members of the system is very striking. This was established in the case of certain dispersed clusters and groups of galaxies (of the cluster in Virgo, Hercules, a group of galaxies in Sculptor, Leonis, etc.). On the other hand, according to Zwicky, large spherical clusters show no expansion signs.

For the complete solution of this problem it is necessary to obtain the largest possible number of radial velocities in several of the closest large spherical clusters.

III. The Principal Facts Concerning the Nature of Galaxies and Their Clusters

Observations show that the forms and internal characteristics of the galaxies are extremely diverse. In order to determine the nature of galaxies, it is extremely necessary to have a sufficiently complete and simple system of the classification of galaxies. It is definitely evident that the more profound physical meaning that the criteria /15 used as a basis of this classification have, the more useful it will be for the solution of the questions of extragalactic astronomy.

The classification of Hubble, the most widely accepted of our time, is founded on the study of the external form of observed galaxies. It has proved to be exceptionally useful in relation to the overwhelming majority of galaxies, and all of our present information is reduced to the data concerning the external forms, the intrinsic brilliance, and the visible diameter. The latter two parameters are not in themselves characteristic of the system as long as the distance is unknown. However, within recent years, it has become possible to estimate the absolute brightness and linear diameter of a very large number of galaxies included into rich clusters inasmuch as it became known that the brightest members of these clusters are always super giants, the absolute magnitudes of which are approximately $-21^{m} \cdot 0$. Comparing this absolute magnitude with the visible magnitude of the

brightest members, we can roughly evaluate the distance, as well as the luminosity and absolute dimensions of all remaining members. As was indicated in the beginning of this report, the range of luminosity of the galaxies in clusters is very large. Gradually, it became clear that the class of the luminosity of a given galaxy (super giants, objects of moderate luminosity, dwarfs, or objects of extremely low luminosity of the type of object of Zwicky in Capricorn) in many cases has a greater significance than even its form. Let us mention once again that the super giant galaxy contains tens of millions more stars than any galaxy of extremely low luminosity.

For the understanding of the characteristics of a galaxy, the study of the nature of its central region and, in particular, the problem of the presence of a central nucleus which is small in magnitude has great significance. It is desirable that new attempts to construct a classification should consider the significance of luminosity, and also, that the assignment of class should determine the role of the central parts and, if possible, the nucleus itself. Finally, it is possible that there are other still unknown parameters which are extremely important for the description of the state of galaxy.

Recently, the proposed classification by Morgan, taking into consideration the concentration of the luminosity to a known degree, answers one of these hopes. However, the assignment of Morgan's class leaves the luminosity undetermined. In the recent works of van den Bergh an attempt was made to introduce a parameter from an observed form of the galaxy, but in essence determining its luminosity. This is a very successful principle. Unfortunately, however, the classification of van den Bergh is not universal and encompasses only the spirals of the latest types. It is, therefore, necessary to assume that, in the future, new classifications will be proposed. These will pose the question of determining the essential parameters of each galaxy.

The most important accomplishment of the second quarter of /16
our century was the introduction of the existence of subsystems in galaxies (Lindblad, Kukarkin, and Baade) and diverse types of stellar population. In certain galaxies, for example in the systems of the EO type, we have a sufficiently large homogeneity of population. In such cases it can be stated that the entire galaxy is composed of only one subsystem. This is particularly true in relation to such members of the Local Group as the system in Sculptor and the galaxies M 32 and NGC 147. Evidently, contrary to the opinion of Baade, we do not consider all such systems to be composed entirely of type I star population (the population of spiral arms). However, in many cases, these galaxies are, in fact, superpositions of two or more subsystems containing diverse types of population.

Thus, the lens-like galaxies (SO) are composed of two subsystems which are, in turn, composed of a stellar population of a globular component and a disc. Giant spirals of the M 31 type are composed of a globular component, a disc, and spiral arms. It is possible that a more detailed division is necessary. For us, however, it is important that, in the given case, various subsystems are superposed.

Available data indicate that the populations of various subsystems travel diverse independent paths of evolution.

There is a basis for considering that the average age of stars of various subsystems is also diverse. It is accepted that, if the dynamic interaction is not considered, each of the subsystems lives its own individual life. This is particularly important because of the description of galaxies as component systems having been derived as the result of the simple superposition of subsystems.

Concerning the relative independence of diverse subsystems entering into the composition of one and the same galaxy, it has been asserted that the degree of development of one of the subsystems (in a sense of the richness of the subsystem and its dimension) is not dependent on the degree of development of the other subsystem.

Thus, for example, the globular subsystem of the galaxy M 31, by its density and dimensions, does not distinguish itself from the normal galaxy of the EO type, which possesses an absolute magnitude of about $-19^{m} \cdot 0$. Meanwhile, the latter does not contain the population of a flat subsystem and spiral arms, while M 31 has conspicuous spiral arms and a richly populated disc.

Holding this point of view, those systems which occupy an intermediate position, i.e., those in which one of the subsystems is developed very firmly while the other is comparatively poor, are also of interest. A remarkable example of this is NGC 5128 (radio source /17
Centaurus A) which, on overexposed photographs, becomes a giant elliptical galaxy. However, in its central part, it contains a weakly developed flat subsystem into which much of the absorbent substance goes. As was shown in the investigations of Berbidge and wife, based on the measurements of radial velocities in this flat subsystem, the equatorial plane of the latter is approximately perpendicular to the equatorial plane of the elliptical subsystem. This is a good illustration which confirms the independence of the subsystem. Another interesting example is the galaxy NGC 3718. The spiral arms of this galaxy possess little prominence; however, unlike NGC 5128, they extend far beyond the limits of the volume occupied by the globular subsystem.

In this galaxy, the plane of the concentration of the dark matter is inclined toward the equatorial plane of the elliptical subsystem by approximately 25 degrees. This also speaks for the independence of the subsystem.

It may have been possible to bring in contrary examples when the spherical subsystem was weakly developed and the flat one was firmly represented. It is evident that the great Cloud of Magellan (Nebecula Major) can serve as such an example. There is a spherical subsystem in this cloud. It therefore follows, if only from its presence, that this subsystem has at least 30 globular clusters similar to the globular clusters in our galaxy and in M 31. Unfortunately, the other objects of spherical subsystems are difficult to distinguish when the population is of a flat nature. Therefore, it is difficult to say that the spherical component of the great Magellanic cloud is similar to elliptical systems. Judging by the distribution of globular clusters and by their number, this should be an elliptical galaxy of a moderate luminosity ($M \sim -16^{m}$) possessing a low density gradient from the center toward the border. It is known that, during the transition from super giant elliptical galaxies to elliptical galaxies of a moderate and low luminosity, there occur more and more often objects indicating a low density gradient.

We spoke previously about the comparatively independent and diverse subsystems entering into one and the same galaxy. However, in one respect, the relation between subsystems is almost constantly observed rigorously. We have in mind the presence of a common center. The center of the globular subsystems is coincident with the center of the disc and, in addition, with the region from which the spiral arms emerge. As is known from the observations of the neighboring galaxies of high luminosity, the nucleus is normally situated in the center and has dimensions of several parsecs (less than the diameter of the common globular cluster). Naturally the thought arises that the origins of individual, almost independent, subsystems are in some way bound by the presence of the indicated nucleus.

In certain galaxies, no traces of nuclei have been discovered. /18
This is the case of NGC 185 or in the case of the system in Sculptor. However, we will turn our attention to the absolute magnitude of the nuclei which we have examined. The photographic magnitude of the nucleus of M 31 is equal to $-11^{m} \cdot 6$. In M 32, it is equal to $-11^{m} \cdot 1$. In M 33, we have $M_{pgn} = -10^{m} \cdot 3$. In NGC 147, we have $M_{pgn} = -5^{m} \cdot 0$. The impression is created here that the absolute magnitude of the nucleus decreases with the decrease of the density gradient. Therefore, it should be expected that in NGC 185, and in the systems of the Sculptor type such as may be in the Magellanic clouds, the nucleus should have an even

lower luminosity than in NGC 147. If it is on the order of $M_{pg} = -2^m$, then it is apparent that the nucleus will become lost among the stars. Let us note that in the Magellanic clouds the nuclei will be imperceptible even if they have $M_{pg} = -5^m$. Therefore, it is premature to make a conclusive deduction about the absence of nuclei in these systems. However, if the nuclei exist in them, then they should possess little prominence.

It was previously indicated that the concentration of subsystems in each galaxy is strictly observed. However, there are individual cases of the break-up of the concentration. NGC 4438 in the cluster of Virgo is a possible example where two subsystems are clearly shifted relative to one another.

There is a certain similarity between galaxies and clusters of galaxies. This is expressed in the fact that, just as with the galaxies, it is possible to divide the stellar population into two basic types. It is also possible to attribute the numbers of clusters of galaxies to two diverse types of population. To the first type belong the spiral and irregular galaxies; to the second, the elliptical and lens-like (SO) galaxies.

The rich, spherical clusters of galaxies of the type of cluster in Coma Berenices contain primarily a type II population. The dispersed clouds of galaxies, similar to the nearby cloud in Ursa Major, practically contain no elliptical galaxies of high luminosity. The close-by group of galaxies in Sculptor ($m-M = 27^m \cdot 0$), investigated by de Vaucouleurs, contains not only elliptical galaxies but also galaxies of the SO, Sa, and Sb types. This group contains only the spirals of the last subclass. The scattered cluster in Virgo contains spiral systems as well as giant elliptical systems.

The question is whether or not it is possible in this case to talk about the superposition of diverse systems in one cluster. It is necessary to recognize that not in all cases are signs of the unification of two quasi-independent subclusters observed in one cluster. However, in certain cases, there is very clear evidence in favor of this. Thus, in the cluster of Coma Berenices, one of the central galaxies (NGC 4874), which is a super giant of the SO type, is clearly surrounded by a symmetrical cloud of elliptical galaxies of lesser luminosity. Superficially, this group is very similar to the galaxy NGC 4486 surrounded by globular clusters. Only in this case, the spherical clusters are substituted by elliptical galaxies of moderate luminosity. And thus, this group of elliptical galaxies, with NGC 4874 in the center, are somehow laid onto a rich cluster possessing a lesser density gradient. /19

Evidently, in the case of dispersed clusters of galaxies, we can find more phenomena attesting to the superposition of individual groups. A very good example of this is the chain of bright galaxies M 84, M 86, NGC 4435, NGC 4438 and others in the cluster of Virgo. As Markaryan indicated several years ago, this chain is not a random formation but was placed on the cluster in Virgo as a certain independent group.

It is entirely possible that the dispersed clusters of galaxies generally are the result of the compilation of certain similar groups which results in their irregular form.

In this relationship, it is necessary to recall the existence of clusters (or groups) which are composed of one central galaxy surrounded by a lesser or greater number of objects of a low luminosity. The group around M 101 is attributed to a number of similar objects. We underline this fact because, in these cases, the central galaxy and its weak satellites undoubtedly have a common origin. However, it is necessary to mention that, along with these systems, there are groups composed almost exclusively of super giants. The quintet Stephan is an example of such a group. Around these super giants, contrary to the previous case, we do not observe any noticeable number of galaxies of low luminosity. However, it is not impossible that a sharp break exists in the function of luminosity and this system contains a certain number of galaxies with absolute magnitudes lower than the specific magnitude which can still be observed. The facts here, combined with those mentioned in the beginning of this report concerning the exclusive state of the position M 31 and our galaxy in the Local Group, show the great cosmogonical significance of the super giant galaxies in groups and clusters.

From the above said, it is also clear that, together with the research on rich clusters of galaxies, it is extremely important to have as much data as possible concerning the comparatively poor groups and, in particular to clarify the possibility of the existence of isolated groups, which are composed exclusively of galaxies of low luminosity. If there are no such groups, then this would mean that, in the /20 formation of dwarf galaxies, the cosmogonical processes occuring in galaxies of high luminosity play a decisive role.

In spite of certain successes in the studies of the nature of stellar populations of galaxies and diverse subsystems, it is still necessary to acknowledge the fact that merely the first steps have been taken in this direction.

It is necessary to have further accumulation of data concerning the composition and population on the basis of spectroscopic data (in the direction indicated by Morgan and Mayall) and the quantitative analysis of spectro-photometric curves (Markaryan and others).

Another important problem is included in the analysis of the nature of galactic arms. With one and the same degree of exposure and length of the arms, the relative richness of their associations is definitely diverse in various cases. By finding a correlation between the nature of the arms and the other parameters of the galaxy, an approach will be made to an understanding of the causes of the indicated diversities.

Of especially great interest are the spiral galaxies with crosspieces (SB). Unfortunately, we do not appreciate fully the difference between the populations of the crosspieces and the arms. It is known only that, generally, the color of the crosspieces is significantly redder than the color of the arms, and that the arms therefore contain a relatively high quantity of young stars. It is especially important to clarify that the crosspieces are rich in clusters and in stellar super giants.

IV. The Broadened Understanding of the Phenomena of Superposition

Individual cases were previously mentioned where the centers of subsystems comprising the given galaxy were shifted with respect to one another. However, we know of other galaxies which are double galaxies, but in effect are bound together by a material medium, and therefore can also be considered as solitary systems. Some good examples are M 51 and the galaxy NGC 7752 - 7753. It is natural in this case to assume that we have an example of a system in which the centers of the subsystem separated. Another good example is IC 1613, where along one side of the main mass of the galaxy is located a giant conglomerate of hot stars, a unique super-association which can be considered a part of the fundamental galaxy and a separate galaxy-satellite. It is highly probable that this super association composed of hot giants was formed much later than the remaining galaxy.*

In connection with this, the concept was evolved that the /21 development of the galaxy is connected with the consecutive formation of diverse subsystems, whereby one or another of these subsystems and sometimes a group of subsystems, also with a new center, might become a satellite of the basic galaxy. This permits us to consider that the

---

*We have such a case in the galaxy IC 2574. A bright super association is located to the North of the main part of this galaxy. They are barely united with each other by an observed arm.

formation of a satellite and the arrival of a new subsystem in the confines of the given galaxy are cognate phenomena.

Furthermore, it is possible to think that these phenomena sometimes accompany one another. As an example, in those cases where the spiral arm unites the center of a given galaxy with a satellite, it is natural to consider that the formation of the spiral arm and the formation of the satellite accompany one another.

Finally, any kind of satellite of the Sculptor type, revolving about the main galaxy, is barely distinguishable from the globular cluster in dimensions and composition of the population. The globular clusters undoubtedly form as a result of inner processes produced in the main galaxy. It is possible to grant the same thing with respect to satellites of the Sculptor type.

V. The Phenomena of Non-Stationarity in Galaxies

Heretofore, galaxies were considered as static formations. However, in galaxies, especially in super giants, phenomena of non-stationary character occur which are of great interest.

We are not concerned here with the processes of the stellar formations in O- and T- associations, although they also have an essential significance for the life of galaxies. We have in mind quicker changes which are directly observable. It is interesting that the majority of these non-stationary phenomena is connected with the nuclei of galaxies and may even be regarded as a manifestation of the activity of these nuclei.

a) From the central region of our galaxy, neutral hydrogen is known to emanate. This phenomenon was observed by Dutch astronomers during radio observations of 21-centimeter wavelength regions. A similar phenomena of gas emission from the nucleus of M 31 was discovered by Munch as a result of the investigation of the line λ 3727. In both cases the intensity of emission reaches the order of one solar mass per year. This result, somehow, does not correspond to the existing values of the masses of galactic nuclei (on the order of $10^7 M_\odot$).

b) Certain galaxies possessing nuclei of high luminosity, as Seifert showed, have an emission line λ 3727 strongly expanded, which corresponds to velocities of motion on the order of several thousand kilometers per second. These velocities exceed the normal escape velocities for galaxies. We deal, therefore, undoubtedly with powerful fluxes of matter, escaping from the nucleus with such great velocities and dispersing

in intergalactic space. Evidently, in this case the quantity of ejected matter far exceeds the corresponding magnitude of our galaxy and of M 31. We can expect that those from the bluish galaxy Aro have an analogous nature. The emission lines of these galaxies are intense in the region around the nucleus.

c) In the very center of the radio galaxy NGC 4486 we have also observed the line λ 3727 and apparently a sufficiently strong flow of gas with a velocity of about 500 kilometers per second. Comparing this with the presence of a radial jet emanating from the center of this galaxy to the surface, and containing plasmas giving intensive radio emission, we come to the conclusion that these plasmas were thrown out from the central nucleus of the galaxy with great velocities. The polarization of light by these plasmas indicates the presence in them of electrons of high energy. However, these plasmas are not formations on the scale of that of the Crab Nebula. The energy of their radio emission measured in absolute units is tens of millions of times greater. If we take into consideration that the extent of radio emission in this case must be at least one thousand times greater, then we come to the conclusion that the supply of energy contained in these plasmas exceeds the total energy contained in the Crab Nebula by one billion times. In other words, these plasmas, as far as their own energy equivalent in the mass is concerned, must be objects on the scale of small galaxies which are found in the photographic rays with respect to their absolute value.

Were these particles ejected from the nucleus of the galaxy in the form of a cloud of relativistic electrons; or, what is more probable were objects expelled from the nucleus continuously creating new fluxes of such electrons? However, it is important that such enormous plasma can be ejected from the nucleus of a giant galaxy, which is related very little to our evidence concerning the masses of the nuclei of galaxies.

d) What occurs in other radio galaxies is far more difficult to explain. We know, however, that the galaxy NGC 1275 (Perseus A) is one of a number of the galaxies of Seifert in which the line λ 3727, observed in the central region, is strongly expanded. In other words, an intense flow of matter from the nucleus occurs in this case. In relation to the information previously developed, the presence of two nuclei in the radio galaxy Cygnus A indicates that the process of separation of the nuclei occurred recently. This must lead to the formation of subsystems with diverse centers and, in the future, to the formation of double galaxies.

In any case the example NGC 5128 (Centaur A) also supports the fact that the nuclei of the galaxies are capable of emitting either /23
a huge cloud of relativistic electrons or some substance capable in the future of creating such clouds.

In one way or another, the radio galaxies appear to be systems in which the central nuclei manifest tremendous activity including the creation of new plasma, new subsystems and possibly new galaxies. Therefore, in the given case, we can freely speak of the cosmogonical activity of the nuclei, even though they are unknown to us with respect to the type of matter that this activity manifests.

e) We are aware of giant galaxies from whose central clouds jet streams flow. Among these are bluish galaxies with absolute magnitudes on the order of -15^m, i.e., having a greater luminosity than the plasma in NGC 4486. Some examples of similar galaxies are NGC 3561 and IC 1182. The expulsion of such plasmas is one more aspect of cosmogonical activity in the nuclei of galaxies.

f) The fact that the spiral arms begin in the nuclei of the galaxies themselves testifies to the fact that the origin of the spiral arms is directly connected with that of the nucleus.

g) The radio observations of the center of our galaxy, conducted by Pariskiy and by others, testify to the fact that the composition of the nucleus consisting chiefly of type II stars, sharply distinguishes itself from the composition of other similar stars (for example, globular clusters). The nucleus itself of our galaxy is the source of thermal and radio emissions, while the surrounding region out to a diameter on the order of 500 parsecs is the region of strong non-thermal radiation. This points to the fact that the physical composition of these indicated nuclei differs strongly from the state of the normal stellar groupings.

One of the most important problems with which we are confronted in the area of the study of the flux of matter and ejections from galactic nuclei is a transition to quantitative evaluations of the ejected masses. This also pertains to galaxies whose central regions emit lines of radiation as well as to radio galaxies and to other cases where we have a case of discrete ejections. These scanty facts that we already have show that these data can lead to contradiction in the law of the conservation of energy (and matter) which is limited in its present state by the forms of energy which we know and require for generalizations of this law.

CONCLUSION

Thus, the major processes in the life of large galaxies is determined by the activity of their nuclei. This activity is expressed in various forms which were mentioned earlier. Of extreme interest, however, are two forms of nuclear activity. One of them is related to the /24 formation of spiral arms and the other with the formation of stars and

stellar clusters of the spherical component. Apparently, these phenomena occur in various stages of development and are accompanied by corresponding changes of the nuclei. In addition, it is necessary to note that the creative process itself of each type of subsystem in various cases should have a distinct character. Thus, for example, the galaxy M 32 obviously does not contain globular clusters while other satellites of the Andromeda Nebula, NGC 205, contains at least 9 spherical clusters. Most astonishing is the fact that globular clusters occur in galaxies with very low density gradients. If the hypothesis of the formation of galaxies from the initial diffuse clouds is accepted, then it seems natural that such dense formations such as globular clusters would arise in systems where there are regions of very high density, i.e., where there are also high density gradients. Of course, such qualitative judgements cannot be considered as sufficient. It is only necessary that the share of globular clusters per unit necessary for luminosity of the spherical population changes from system to system. Thus, we obtain the additional parameter for the characteristically spherical systems and subsystems. That this parameter is connected with other parameters of the same systems (total luminosity, gradient of density) must be clarified by observations.

Statistical data concerning multiple galaxies and clusters of galaxies reveal the fact that these systems could not be formed by mutual capture previous to the independent galaxies. Therefore, it is necessary for the components of the indicated systems to be attributed to a common origin. This problem was discussed in detail in our report to the Solvay Conference in 1958.

In the light of the data mentioned above, concerning the ejection of the nuclei of the coagula transforming them into complete galaxies of moderate or high luminosity, and concerning the division of the nuclei, an understanding of the origin of multiple systems and full groups becomes probable as being the result of the division of one initial nucleus into several nuclei. It is possible that this division occurs sequentially.

In these cases, when there is a central galaxy of high luminosity in a group, the formation of weak galaxies must be connected primarily with the activity of the nucleus of the galaxy of high luminosity.

The radio galaxies often appear to be some of the brightest members of the clusters in which they appear. This is explained by the very high activity of the nuclei of the galactic super giants. If on the other hand, in a cluster there exists one clearly dominating galaxy, it is normally a radio galaxy itself.

The observations show that, although all the large clusters /25
contain super giant galaxies, only a small part of the latter are radio

galaxies. In this way, the radiation activity must be a relatively short phase in the history of the development of galaxies. It should be assumed that the separation of radio-emitting agents is the phenomenon which accompanies the expulsion from the nuclei of more powerful masses. This possibly occurs only in the definite state of that or another cosmogonical process.

Although extragalactic astronomy has great possibilities with respect to the study of nuclear activity, all of the evidence concerning various aspects of this activity is extremely scant. Still less is known about the parameters characterizing the integral properties of these nuclei (luminosity, mass, color, dimensions, rotation). Finally, nothing is known about the inner structure of these nuclei. The most extensive field of research occurs in connection with this region of extragalactic astronomy. For example, the following problems arise:

1. Do all galaxies have nuclei; if not, what are the characteristics of galaxies which have no nuclei?

2. The determination of the integral characteristics of the nuclei for a majority of galaxies. Thereby it is necessary to take into consideration the difficulty of this problem relative to the galaxies with high gradients of density. In addition, it should be mentioned that many galaxies of the type Sc have a nucleus which is so well distinguished that it can be investigated without much interference of the near-nuclear central plasma.

3. The determination of the dependence between the integral parameters of the nuclei and the integral parameters of the galaxy.

4. Investigation of the spectrum of nuclei for clarification of emission lines, phenomena of rotation, and flux.

5. The investigation of the bond between the nucleus and the crosspiece in galaxies containing the crosspiece. The relation between the crosspiece and the phenomenon of flux from the nucleus.

6. The research on galaxies with multiple nuclei. The study of radial velocities of individual components of such nuclei.

7. The dependence of the number of globular clusters on the nature of the nucleus of the galaxy.

Although, we previously mentioned certain evidence of a cosmogonical character, pertaining to the origin of the galaxies, we constantly tried to remain on the basis of facts and not become involved in remote speculation. The analysis of observations shows that the

phenomena pertaining to the origin of galaxies are so unusual that it would be impossible to imagine them arising out of some theoretically preconceived situation. Here again, we are confronted with the astounding phenomenon constantly repeating itself in the history of /26 science: when it intrudes into a new region of phenomena, it finds unexpected new rules quantitatively going out beyond the limits of previous concepts. This makes each such region of phenomena so much more interesting. Therefore, it is necessary for us to collect even more carefully the facts and observations for the augmentation of factual data and more exact evidence concerning actual objects. We must gather much information concerning the structure of the various parts of the galaxies and make a careful analysis of this evidence. This might help us in the solution of difficult problems which occur here.

ON SOME ASPECTS OF AMBARTSUMYAN'S HYPOTHESIS ON THE ORIGIN OF GALAXIES /27

By

B. A. Vorontsov-Velyaminov

Facts and considerations in favor of the different aspects of Ambartsumyan's hypothesis and the difficulties encountered are discussed.

The opinions of V. A. Ambartsumyan concerning the origin of galaxies often, especially abroad, interpret the hypothesis concerning unstable multiple systems and clusters of galaxies. Actually his concept is much broader than the hypothesis on the mechanical instability of groups of galaxies. The assertions that the nuclei of galaxies are super dense formations and are dynamic were essential arguments for this hypothesis. Their separation generates multiple galaxies, and the ejection of smaller masses of a super dense, pre-stellar substance leads to its transformation into spiral branches, globular stellar clusters, and diffuse matter.

In connection with some of these conditions, it appears to be expedient for us to share a number of concepts.

Page (Ref. 1) advanced an interesting idea for a statistical verification of the fact whether double galaxies originated concurrently, in particular in the same manner of division of the single type, or whether they originated by capture. The preliminary results which he obtained led to unexpected and improbable deductions. It is quite clear that his idea was at first quite simple, but became complicated in practice by a number of circumstances which must be studied.

We are confronted with the fact that the concurrent physical origin of double and multiple systems definitely exists. This was proved to a sufficient degree by Ambartsumyan and others; in particular, interacting vapors whose physical dualities were so frequently evident that they could not have been caused by random encounters of galaxies.

The theory of capture in connection with the cosmogonical /28 hypothesis of Schmidt was studied very circumstantially. It was shown that the specific approach of three bodies is necessary for capture to occur. Therefore, the probability of capture among isolated galactic fields in which the interacting galaxies are numerous is very low.

Variation of velocities in physical vapors, on the average, is also smaller than the characteristic velocities of the galaxies. As we see it, this rigorous proof supplements the critique of the theory of capture.

Relative to the question concerning the instability of clusters in groups, and to their positive energy, the answer is less definite. In the majority of cases, authors are inclined to draw conclusions concerning the instability of the small groups which they have studied, using the data concerning radial velocities. However, the calculations include additional suppositions and inaccurate data. The assumed distance to the groups, the suppositions concerning the chaotic equally probable distribution of directions for combining a small number of galaxies in pairs, etc., enters in here. An analysis of the possible mistakes in the calculations of definitions by Limber and Mathews (Ref. 2) shows significant unreliability in the results. It is interesting to note that, accidentally or not, the determination of apparent group components' radial velocities lead to values differing by hundreds and thousands of kilometers from the mean velocity of the group. This has occurred in several cases. The first point of view, concerning whether or not they belonged to a group, became doubtful in spite of the visual size and position of these galaxies in the sky. Accepting them as members of a group leads to the deduction that the formation of the group had a definitely explosive character. We assume that the absence of visible turbulence in the structure of such components is no proof of their random projection, since evidently distortions of forms in the structure have a non-gravitational nature and do not always manifest themselves in previously known small spaces, and are sometimes perceptible in large distances between components. In the case of clusters of galaxies the problem is even more complicated since it is more difficult here to distinguish the members of the cluster. The dimensions of the clusters are always significantly less than the distances to them. A background galaxy remote from the center of the cluster, let us say at 5 radii of the cluster along the line of sight, can not be distinguished from the other members of the cluster. We find confirmation of this opinion, for example, in the fact that many investigators already do not consider that the cluster in Virgo is one cluster.

To all of this, the additional suspicion was voiced by Holmbert that, in the determination of the radial velocities, systematic errors exist, depending on the brightness of the nucleus of the galaxies.

In general, however, it is more probable that, at least in the case of multiple galaxies, we have the matter of their mutual recession.

Ambartsumyan considers the systems of the Trapezium type as /29 unstable and still young. Berbidge (Ref. 3) proposes the same thing with

respect to the compact chains of galaxies which I have already discovered. However, among these groups, there are elliptical galaxies which are considered to be very old formations. However, I indicated (Ref. 4) that the existence of very long, uniting fibers and shafts speaks in favor of their prolonged existence and significant stability. With respect to the elliptical galaxies entering into the chains, and in the Trapezia, the following suppositions are possible:

1) These elliptical galaxies are young. However, some of the distinctions from old elliptical galaxies are not known.

2) These are not elliptical galaxies; they merely appear to be such. So far, there are no data in favor of this.

3) The systems into which they go, despite the mechanical consideration, are nevertheless stable and old. Can it not be that a mutual repulsion began arising in the old system, which was previously the stable system, in connection with the evolution of galaxies?

Many are interested in the question of where the astounding energies come from that are necessary for the dispersion of the parts of the divided nucleus of the galaxy having a velocity of hundreds of kilometers per second. If we wish to ignore this question, relating it to the unusual characteristics of super dense stellar matter, then the division of galactic nuclei can be accepted. However, it is difficult to imagine how a mutual recession can occur in already formed galaxies that penetrate one another and have rotations independent of one another.

If we rely on the law of gravity, the existing data scarcely admit the huge masses near the nuclei which, in certain galaxies, is still composed of super dense matter. If, in younger galaxies, such nuclei still exist and display activity, this should be disclosed by observations. Such a sign of activity is considered by Ambartsumyan to be the "ejection" from the nucleus NGC 4486; that this is actually an "ejection" (as everyone calls it) was not confirmed by anyone. It would be necessary to make a more detailed study of the nuclei of galaxies in search of traces of their activity.

A number of new data can be considered as confirming Ambartsumyan's idea concerning the specific role of nuclei although the discovery of the emission of gas from the nucleus of our galaxy and others is difficult to explain by normal phenomena. The unusually rapid rotation of the nuclei is sharply distinct from the rotation of their surroundings, and the presence of dark fibers leaving the very center of M 31 also testifies to the unique nature of the nucleus.

The latter confirms Ambartsumyan's theory that the spiral branches, formed by the injection of super dense bodies from the nucleus, are /30 not very massive. On the other hand, the formation of spiral branches formed in this way is difficult to imagine. In the first place, a fragment of the super dense substance being ejected must have moved along a straight line. In the second place, for the formation of a spiral branch, it should have continuously separated matter from itself along the trajectory while transforming into stars and gas. It is true that there are galaxies of the type Sc which have spiral branches composed primarily of remote globules, but such cases among the spirals are generally rare. It is difficult to explain the fact that the spirals have two predominant branches, that often the branches start from the ends of crosspieces, and not from the nucleus; that there are galaxies with most of their branches leaving the nucleus in one equatorial plane. There is the impression that the spiral branches had developed inside the already existing disc, and not that the discs were a result of the dissipation of the spiral branches. These branches are often definitely not bound to the nucleus.

The branches frequently begin on the periphery of the disc, or within it. Sometimes there are inner and outer spirals not connected to one another. They sometimes leave the ring in two spirals from one point, or are completely separate from the inner parts.

Ambartsumyan's admission that in the presence of the formation of various elements in the structure of the nucleus, galaxies eject fragments in various planes. This can be confirmed by various examples to a certain degree. Not only do structures occur having various orientations of the major axis in one and the same plane, but also, in certain systems, for example, M 82, the fluxes of matter are directed perpendicular to the main plane of symmetry.

It seems to us that the concept according to which, in multiple galaxies, the components originate concurrently in immediate proximity to one another and diverge in all directions in the process of their own formation is more probable. There sometimes arises a repulsion of a certain type among them which causes their mutual recession regardless of gravity. The substance which is subjected to the greatest repulsion forms shafts. The great tenacity of stellar systems, which must be accepted, causes the formation of crosspieces. The stability of the shafts and of the crosspieces is the same as the stability of the spiral branches, and is significant. The close proximity of the galaxies somehow impedes the development of the spiral structure and causes the phenomena of "destruction of the facade". These phenomena of repulsion in the interacting galaxies must be akin to the phenomena causing the mutual recession of the galaxies in the group.

It should be recalled that the Metagalaxy, although composed of galaxies is, at the same time, an original continuous medium in which the galaxies are brought together. Experimental physics has never had dealings with anything similar. Therefore, in the world of galaxies, /31 it is possible to state that the phenomena will be completely unexpected, and so far not understood.

We are ready to admit that pairs of galaxies have quite different characteristics other than the attraction of the total mass, just as the characteristics of molecules are distinguished from the characteristics of the components of atoms.

Certain pairs of living organisms are capable of self-reproduction. Certain pairs of combining atoms possess new characteristics, and it is impossible to derive all of these characteristics from a knowledge of the characteristics of individual components. Something similar may occur in the universe of galaxies. Confirmation of this can be seen in the fact that, according to Zwicky, the clusters of galaxies do not interact gravitationally with each other.

REFERENCES

1. Th. Page. Astronomical Journal, 61, 1961.
2. N. Limber, W. Mathews. Astrophysical Journal, 132, 286, 1960.
3. G. Burbidge, M. Burbidge. Astrophys. J., 131, 742, 1960.
4. B. Vorontsov-Velyaminov. Astronomicheskiy Zhurnal, 34, 8, 1957.

INVESTIGATION OF DISTANCES, MOTIONS AND DISTRIBUTION OF GALAXIES IN A SPHERE WITH A RADIUS OF 15 MEGAPARSECS

By

Yu. P. Pskovskiy

The utilization of galaxy classifications, as proposed by van den Bergh, and the more precise definition of mean absolute magnitudes of various type galaxies improves the method of integral values. The zero-point of the relation period-luminescence may be determined accurately only according to stars, which have individual absolute magnitude evaluations (Cepheids in dispersed clusters and spectral parallaxes). The modulus of distance M 31, which is determined by Cepheids in the presence of a new zero-point, and the relation period-luminescence of the Cepheids according to Arp is found to be equal to $23^{m}.95$. This agrees well with evaluations made in respect to novae, if we take into consideration the fact that the novae are weakened by absorption within the M 31. /32

The spatial distribution of galaxies within an area with a radius of 15 megaparsecs is considered by means of simple diagrams. This distribution confirms the existence of a galactic cloud with a central concentration in the Virgo cluster (hypergalaxy). The Fornax Cluster is not a hypergalaxy, but is only an arm of our hypergalaxy.

The character of motion in the hypergalaxy was investigated according to the red shift of 370 galaxies. The principal values of the tensor of the speed of deformation were found. De Vaucouleurs' model on the rotation of a complex of galaxies around one axis is descriptive, but very rough. There does not exist a single axis of galactic rotation. A galaxy rotates around a central plasma along chaotically oriented orbits. The motions within a hypergalaxy very much resemble that of a cluster.

1. If the distances to the nearby galaxies are correctly evaluated with the help of reliable distance indicators, then distances to the receding galaxies are estimated either according to integral magnitudes or according to the red shift of the galaxies. One must not consider any of the recent methods satisfactory. The integral absolute magnitudes of the galaxies are sufficiently different, and the use of the average absolute magnitude, instead of the individual magnitude, leads to errors in evaluations of distances on the order of one or two times. The law of the red shift, which was established using the distances determined by the method of integral magnitudes has still other /33

sources of errors: the great dispersion of the red shift and the neglect of the dependence of the effect of red shift on direction. It is understood that the red shift, as a means of estimating distances is generally unsuitable. This is realized if a study is made of the regularities of motion of some groups of galaxies which surround us.

In connection with this, it is natural to follow the path of the improvement of the method of integral magnitudes and to continue the search for new forms of distance indicators. Since the dispersion of absolute integral magnitudes of any type of galaxy is less then the general dispersion of all galaxies without the division into subtypes, then by utilizing the average absolute magnitudes of the individual subtypes of galaxies, it is possible to decrease the errors in the method of integral magnitudes (Ref. 1). In irregular and spiral galaxies, van den Bergh (Refs. 2-4) has discovered a correlation between the degree of development of the spiral structure and the main body of the galaxies with their integral absolute magnitudes. Keeping in sight the investigation of the motions of galaxies, we have determined the average integral magnitudes of galaxies in various classes of luminosity (Ref. 5). The data concerning the galaxies of the cluster in Virgo, in near-by groups, and in galaxies in which the "super-novae" were observed, serve as the material for such determinations.*

In this way, from the majority of subtypes of galaxies, we have /34 obtained more accurate evaluations for the values for the average integral absolute magnitudes.

2. Within the last few years some works appeared that are devoted to the most reliable distance indicators, i.e., Cepheides and novae (Refs. 9-18). In the improvement of the material of observations arise some contradictions in the values of modulus of the distance up to M 31, which is the basis for the remaining extragalactic distances.

---

*Lundmark (Ref. 6) was the first to turn his attention on the super-novae as indicators of distance. The theoretical considerations (Ref. 7) and the comparison of the material in the observations on super-novae, which have flared up in the groups and clusters of galaxies show that, in galaxies that are weak in luminosity, the brightest super-novae of type I (Ref. 8) flare up. Although the maximal magnitudes of the super-novae are roughly determined by the extrapolation of the curves of luminosity, the estimations of distances by means of these indicators merit serious attention. As for the range in which they allow distance evaluation, they have no equals among the stars.

The study of the problem of the zero-point of Cepheides leads us to the conclusions voiced by the X Congress of the International Astronomical Union by O. A. Melnikov. The values of the zero-point, founded on the statistical evaluations of the average absolute magnitude of Cepheides, give only qualitative results. More accurate results can be expected from the methods which give reliable individual evaluations for absolute magnitudes for Cepheides belonging to clusters and by the luminosity effect. The material concerned with Cepheides belonging to clusters is limited and as for the effect of luminosity, the further accumulation of observation, all data is still not accumulated.

Nevertheless, according to six Cepheides in the dispersed clusters (Refs. 9-14) there is the following relation. The period-luminosity relation in the form determined recently by Arp (Ref. 15) for the Cepheide in the Magellanic Cloud is

$$M_{pg}^{med} = - 0^m.81 - 2^m.25 \lg P \ (0.0 < \lg P < 1.8).$$

Adding this relation to the Cepheides in M 31 investigated by Baade (Ref. 16), we obtain a distance modulus of up to $0^m.30$ less than the modulus of Baade (Table 1) for the distance to M 31 through the Cepheides outside the absorbent layer of the nebula. Incidentally, Sandage (Ref. 18) considers the modulus of the distance to M 31 more reliable when based on the lustre-luminosity curve of the novae at a maximum in the presence of the zero-point of this relation according to Schmidt (Ref. 17).

Table 1

Explorers	Material	Modulus of the Distance to M 31 According to this Paper	Correction on the Absorbence in M 31	Corrected Modulus of the Distance to M 31
Baade (Ref. 16)	Cepheides	$24^m.25$	-	$24^m.25$
Schmidt (Ref. 17)	Novae	24.6	$-0^m.75$	23.85
Our evaluation	Cepheides	23.95	-	23.95
Mayall, Kron (Ref. 19)	Globular clusters	23.5-24.0	-	-

This modulus is $0^m.65$ greater than our evaluation. If we take into consideration that the novae in M 31 are somewhat weakened by the absorbent material in the nebula and that, according to the diagram in Fig. 5 of the work of Baade (Ref. 16), the mean absorption obtained is on the order of $0^m.75$, then the difference in the evaluations of the modulus of the distance to M 31 through the Cepheides and novae can be attributed to the difference in absorption. It is necessary to note that the zero-point of the relation for the novae, the zero-point determined by Schmidt, possibly contains errors owing to the inaccuracy of the methods used regarding the absorption in our galaxy. /35

In our opinion, the evaluation of the distance modulus up to M 31 according to new globular clusters (Ref. 19) and according to the Cepheides, after a consideration of the previously indicated conditions, coincide well. In this example, it is evident that it is necessary to consider the absorption in the galaxies if their distance modulus is being determined through the Cepheides, the novae, and the super giants located in the absorbent layers. Poor calculation leads to systematic errors in the estimates of distances, and in the case of M 31 it leads to mistakes in the whole scale of extragalactic distances, since the distance to it serves as the basis for this scale.

3. With the help of mean absolute integral magnitudes of the galaxies in the sphere with a radius of 5 megaparsecs around our galaxy, the number of elliptical and spiral galaxies was counted. Out of 266 galaxies, 28 percent were elliptical and 72 percent were spiral (Ref. 1). The same correlation is observed in the very large clusters of galaxies of irregular form: in Virgo (Ref. 1) and in Hercules (Ref. 20).* The similarity of these regions of the Metagalaxy is undoubted.

In connection with this, it is possible to mention de Vaucouleurs' hypothesis (Refs. 24-27) and the earlier hypothesis of Shapley and Reiz concerning the fact that our galaxy and its neighbors belong to a supercluster with its center in the cluster in Virgo. By observing this supercluster from within, according to de Vaucouleurs, we see in the sky a "Milky Way" of galaxies. De Vaucouleurs introduced a system of

---

*Calculations of the number of spiral and elliptical galaxies in our environs were conducted earlier by Yu. R. Yefremof (Ref. 21) and also according to Catalog X. H. Shapley and A. Ames (Ref. 22) based their assumption on the fact that the elliptical galaxies of the Metagalactic field were weaker than the spiral ones at $4\text{-}5^m$ (i.e., of the type NGC 205 and the systems in Sculptor). If this were so, the relation on the red shift visible magnitude [see Fig. 3-10 in (Ref. 23)] would be displaced for the elliptical and spiral galaxies relative to one another by $4\text{-}5^m$.

coordinates corresponding to the main plane of this "Milky Way" of galaxies. It made sense to call this system a hypergalactic system of coordinates. Accordingly, we shall call this supercluster the hypergalaxy.* We have checked the visible distribution of elliptical /36 galaxies, concentrated according to de Vaucouleurs (Ref. 25) toward this plane and have confirmed the existence of a hypergalactic plane as a plane of symmetry of their visible distribution. We utilized the system of hypergalactic coordinates of de Vaucouleurs for the study of the distribution and motion of galaxies in the sphere with a radius of 15 megaparsecs.

During distance calculations, we utilized the mean absolute integral magnitudes of the corresponding subtypes and classes of luminosity (Refs. 1, 5, 8), just as in our preceding works we used the physical integral magnitudes according to the published lists of authors (Refs. 22, 23, 30-35) recounted in Holmberg's system (Ref. 30). The absorption of light in the galaxy is assumed to be according to the chart by P. P. Parenago (Ref. 36). The agreement between the method of considering the absorption to be according to the count of weak galaxies squared and the method of the excess of color of the stars (Ref. 37) was confirmed earlier.

4. The spatial distribution of galaxies can be presented graphically if you examine the visible portions of the sphere within a radius of 15 megaparsecs around our galaxy by means of the planes of the hypergalactic equator and the meridian passing through the cluster in Virgo (hypergalactic longitude $L = 105^{\circ}$). The semithickness of the layers in cross section is equal to 2 megaparsecs.

The meridional section (Fig. 1) shows a comparatively homogeneous distribution of galaxies in the southern hemisphere and outside the equatorial zone of the northern hemisphere. In the equatorial zone, is visible a plasma, which is the cluster in Virgo. Traces of the cluster in Fornax $L = 265^{\circ}$ are found near the hypergalactic latitude $B = -40^{\circ}$. A strong elongation of the cluster in Virgo in a radial direction (to our galaxy) is explained by the ficticious dispersion of distances. We have determined distances according to mean absolute integral magnitudes of

*K. F. Ogorodnikov (Ref. 28) calls a "hypergalaxy" that which is universally accepted to be called the Metagalaxy (Ref. 29); and "a Metagalaxy" he calls that part of the "hypergalaxy" which is easily accessible to investigations. In order to avoid confusion, we are using the more common terminology, considering the term supercluster of galaxies to be synonomous with hypergalaxy.

the galaxies, and the deviation of individual absolute magnitudes from the mean is manifested as a scattering in distances, i.e., the radial directions.

While examining the layer near the equator (Fig. 2) we also see a comparatively homogeneous distribution of galaxies in one half of the field (for a sufficiently large part of the volume) and a strong non-homogeneity in the other. The cloud of galaxies is stretched out more than 100° along the longitude and includes within itself our galaxy, its surroundings, and the concentration of galaxies in the southern hemisphere ($210^\circ < L < 270^\circ$). It is interesting to note that the so called "southern supercluster" neighboring ours (Ref. 24) is actually not a supercluster or a hypergalaxy (this is not shown in Fig. 2 since it is situated on the equator at latitude $B = -41^\circ$). This is a comparatively small complex of galaxies having a diameter of about 5 megaparsecs and is not similar to our hypergalaxy by composition of its galaxies. In reality, this complex consists of a cluster of primarily elliptical galaxies in Fornax and /37
of an arm of our hypergalaxy extending to it, which consisted of a mixture of spiral and elliptical galaxies.

In Fig. 2, the ficticious dispersion of the distances was hardly noticeable. However, this dispersion is present: a group of galaxies in Ursa Major ($L = 75^\circ$, $B = 0^\circ$) and other groups in the illustration are extended in radial directions and became one cloud surrounding the central cluster in Virgo. This corresponds to the concept made by de Vaucouleurs and van den Bergh (Ref. 38) concerning the fact that these groups are visible condensations of galaxies in the peripheral part of the hypergalaxy. The hypergalaxy is a visible thickening of the general Metagalactic field without a sharp transition. An analogous picture is also obtained according to red shifts as a measure of distance.

In Table 2, the count of hypergalactic density is shown; i.e., the number of galaxies per cubic megaparsecs. Only the spiral and elliptical galaxies with absolute integral magnitude brighter than -17^m were calculated. The calculations were directed along 30° sectors of the layer near the equator within intervals of 5 megaparsecs. These calculations characterize numerically the distribution of the galaxies in the layer. The greatest density (3.5 galaxies per cubic megaparsecs) is observed in the region of the cluster in Virgo. In the region around the cluster, the galactic density varies from 2.2 to 1.5 galaxies per cubic megaparsec and the peripheral regions from 1.0 to 0.5 galaxies per cubic megaparsecs. The density of the general Metagalactic field on the average is lower than 0.5 galaxy per cubic megaparsec. The density of the arm and the cluster in Fornax is moderate: 1.0 to 0.5 galaxies per cubic megaparsec.

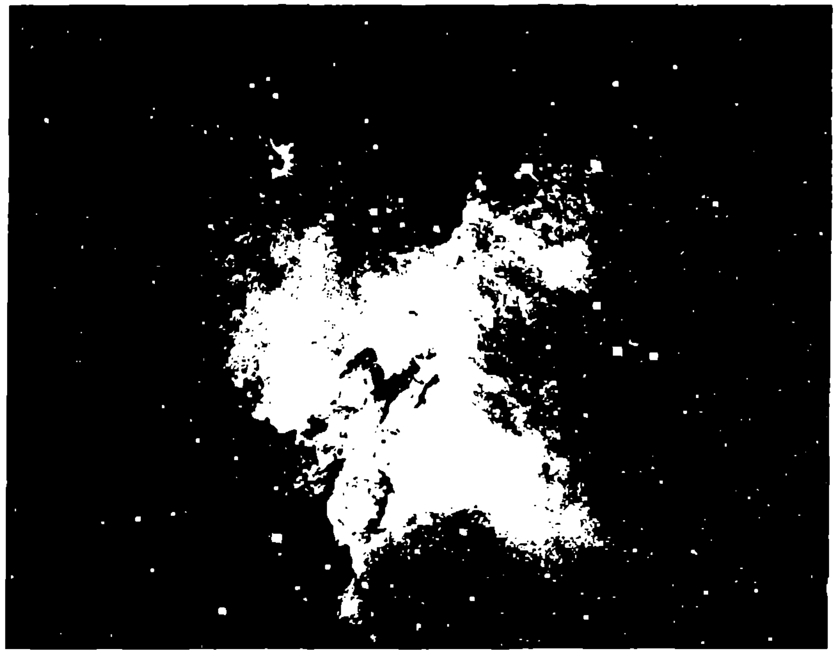


Fig. 1. Diffusing Nebula NGC 6611 with Bright Rings at the Boundary with Dark Matter.

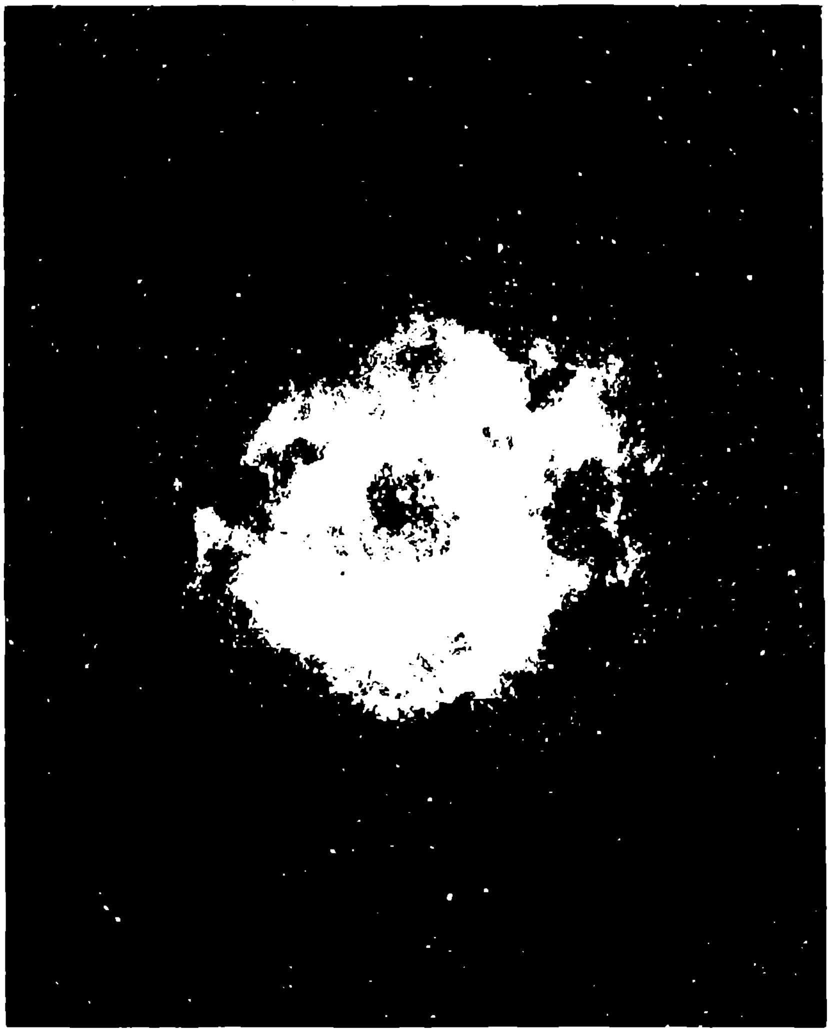


Fig. 2. Diffusing Nebula NCG 2237 with a Concentration of Matter at the Periphery.

Table 2. Density of Bright Galaxies by 1 Mparsecs^3. (Within the Parenthesis are Given the Number of Galaxies with $M < -17^m$)

Layer M parsecs	Sector						
	0—30°	15—45°	30—60°	45—75°	60—90°	75—105°	90—120°
	1	2	3	4	5	6	7
0—5	(7)0,54	(12)0,92	(13)1,00	(21)1,61	(24)2,23	(25)1,92	(23)1,85
5—10	(4)0,10	(28)0,72	(38)0,97	(42)1,08	(69)1,76	(98)2,51	(137)3,52
10—15	(1)—	(14)0,13	(24)0,23	(28)0,27	(20)0,19	(35)0,34	(51)0,49

Layer M parsecs	Sector						
	180—210°	195—225°	210—240°	225—255°	240—270°	255—285°	270—300°
	8	9	10	11	12	13	14
0—5	(8)0,62	(9)0,69	(10)0,77	(11)0,85	(9)0,69	(7)0,54	(2)0,15
5—10	(8)0,20	(18)0,46	(25)0,64	(42)1,08	(34)0,87	(12)0,31	(11)0,27
10—15	(0)—	(0)—	(0)—	(1)—	(1)—	(4)—	(12)0,12

Layer M parsecs	Sector				
	120—150°	105—135°	150—180°	135—165°	165—195°
	15	16	17	18	19
0—5	(10)0,77	(12)0,92	(13)1,00	(18)1,38	(4)0,31
5—10	(60)1,54	(124)3,08	(27)0,69	(49)1,26	(2)—
10—15	(8)0,08	(42)0,40	(0)—	(3)—	(0)—

Layer M parsecs	Sector				
	300—330°	285—315°	330—0°	315—345°	345—15°
	20	21	22	23	24
0—5	(0)—	(1)—	(12)0,92	(9)0,69	(6)0,46
5—10	(10)0,26	(13)0,33	(8)0,20	(12)0,31	(1)—
10—15	(6)0,06	(12)0,12	(8)0,07	(11)0,10	(1)—

5. Many investigators (Refs. 26, 39) have already noticed the so-called "local anisotropy" in our environs of the Metagalaxy. This is shown in the fact that a graph of the dependence of the red shift on distance systematically strays from the linear Hubble law within an interval of 10 megaparsecs [see Fig. 10 in (Ref. 23)]. Besides that, the points on this graph are widely scattered due to the following causes:

1) The distances are not determined according to the method of integral magnitudes; a more accurate definition of distances may decrease this uncertainty.

2) Dispersion of the red shifts. In clusters, for example, the amplitudes of the red shifts reaches 2000-4000 km per second (Ref. 40).

3) The anisotropy of the velocities of deformation speed or in other words, the dependence of the parameter of Hubble on directions L, B. The stronger the band is extended, the sharper is the dependence.

It is possible to properly distinguish the dispersion of the red shifts from the anisotropy of the velocities or the distortion /38 of items either by selecting a model of the motions of the total galaxies being observed, or by determining from observations the parameters of motion of this totality and analyzing the character of this motion according to the magnitudes of these general parameters.

An example of research with the aid of a model is de Vaucouleurs' work concerning the differential (Öort) rotation and the differential /39 expansion (k-effect) of the supercluster of galaxies participating in the isotropic expansion of our region of the Metagalaxy [the law of Hubble (Refs. 26, 27)]. In accordance with de Vaucouleurs, the flattened form of the supercluster indicates the fact that it is in a state of rotation and, according to the formula of Weaver and Troumpler [the generalization of the formulae of the galactic rotation (Ref. 41)], he calculated the parameters of such a motion assuming the center of rotation in the cluster is in Virgo. He compared the model of the parameters calculated in this way with the data of observations, and it was agreed that it was not astonishing since the calculation of the parameters had already determined that result.

A more general investigation of the character of motion of the galaxies in our environs of the Metagalaxy was conducted in 1951 by K. F. Ogorodnikov (Ref. 42) who calculated the components of the tensor of deformation velocities. In his work, K. F. Ogorodnikov proceeded from the hydrodynamic concepts applied to the Metagalaxy. A relativistic

examination of the phenomena accompanying the transmission of light in the Metagalactic medium, which is arbitrary in its characteristics (in other words, in a non-homogeneous anisotropic medium) led A. L. Zelmanov to the expression of the red shift in the first approximation by means of the tensor of the deformation velocities, doppler effect, and the Einstein vector. Due to the low density of the Metagalactic medium, the last effect according to the evaluation of A. L. Zelmanov in a given period of time is insignificant as compared to the doppler effect. Thus, from the relativistic considerations in the first approximation, the expression of the red shift through the tensor of deformation velocities is also found.

From 1958 to 1961 we also calculated the tensor of the deformation velocities by means of the "antipode method"; that is, by the red shifts and distances in the diametrically opposite areas in the sky (Ref. 44), and by means of a method of accounting for components of the tensor of deformation velocities for spherical layers around our galaxy. During calculations according to the first procedure the absorption was obtained in accordance with the cosecant law (Ref. 23) and the distance was obtained according to the method of integral magnitudes. During calculations according to the second procedure only the irregular and spiral galaxies were utilized and the absorption and distances were determined in the manner indicated in Section 3. The components of the tensor were calculated for five layers (Table 3). In each case the main values of the components of the tensor were found. The main semi-axes of a certain ellipsoid corresponding to these values geometrically and also the directions of these semi-axes in space were found. The magnitude of the mean semi-axis was very reliably found. If the supercluster is actually located in the state of Öort's rotation, then this mean semi-axis /41 represents, in itself, the Hubble parameter freed from the effect of the differentiation of the supersystem.

It is natural to compare the parameters which we have obtained with the ones that should have been found in the case of an Öort rotation. As A. L. Zelmanov showed (private announcement) if the hypergalaxy is found in the state of Öort's rotation then the mean semi-axis of the ellipsoid of the deformation velocities is parallel to the axis of rotation of the supersystem, and the bisectors of the angles between the major and the minor axes indicate the direction to the center of rotation. However, any flat parallel motion can be examined as a combination of the successive motion with a rotation about the instant center. If a cluster, or another formation of significant mass, exists in one of the possible directions toward the center of rotations, then it is possible with sufficient basis to assume that actual rotation of the surrounding bodies occurs around this center. Such a state exists in our galaxy: in the direction of the center of rotation, a bright cloud of stars is found in Sagittarius, which is the nucleus of our galaxy (Ref. 28). From

Table 3. Semiaxis of the Ellipsoid of Deformation Speed.

Layer M parsecs	Mean Distance $\bar{r}$, M parsecs	Number of Galaxies	Major Semiaxis D_1, km/sec per M parsec	L_1, B_1	Mean Semiaxis D_2, km/sec per M parsec	L_2, B_2	Minor Semiaxis D_3, km/sec per M parsec	L_3, B_3	Hubble's Constant H=D, km/sec per M parsec	Direction toward Center of Rotation L_o, B_o
"Antipode" Method										
		310	270	227°, −22°	177	260°, −64°	87	142°, +12°	178	94°, +17°
Spherical Layer Method										
4—6	5,25	65	278	**334**, +67	161	342, −23	122	68, + 4	187	111, −35
6,3—7,6	7	67	276	**72**, −46	268	54, +43	189	332, − 9	244	112, −28
7,9—9,6	9	75	244	50, −81	208	34, + 9	179	304, − 2	210	87, −39
10—15	12	87	270	36, +27	221	63, −60	212	130, +11	227	83, +19
15,8—38,0	22	76	244	42, +63	196	32, −26	183	56, − 4	207	139, +29
Total		370						Average	213	106, −11°

K. F. Ogorodnikov (Ref. 42) it has been proved that one of the probable directions toward the center of rotation coincides with a large cluster in Virgo. We will mention that, in this case, the magnitude of the minor axis was used for the parameter of Hubble which, in reality, is the semi-axis; therefore, the direction to the cluster in Virgo was taken incorrectly. According to our results, we have obtained roughly on an average the direction to the cluster in Virgo, although the galaxies of this cluster were excluded from the investigation. This is in favor of de Vaucouleurs' model. However, the direction to the cluster in Virgo, as well as to the center of rotation and the direction of the axis of rotation along each layer, are clearly distinguished. Consequently, the rotation in the hypergalaxy does not occur about a unique axis, as is supposed in de Vaucouleurs' model, or in the Öort rotation, but the galaxies move around a central plasma along orbits which are very distinctly oriented.

The motion of the galaxies in the supercluster reminds us of the motion in an expanding swarm of bees, somewhat compressed into a plane of symmetry of distribution of galaxies. From the point of view of mechanics, this is a problem of many bodies. The motion of galaxies in the first approximation is statistically characterized by six parameters (components of the tensor of deformation velocities, or parameters equivalent to them).

Thus, having proposed the rotation of galaxies about one axis perpendicular to the hypergalactic plane along orbits coplanar to this plane, we come to de Vaucouleurs' model in which, instead of an expanding cluster of galaxies, which are moving in chaotical oriented orbits, Öort's law of the differential rotation with differential expansion /42 in a hypergalactic plane is introduced. The coordinates of the direction of the axis of rotation (two independent parameters), the direction to the center of rotation (one independent parameter), the parameter of rotation (the Öort parameter), and the coefficients of the isotropic and differential expansion of the system correspond identically to the six tensor components. The difference of the real motion obtained according to this model leads to the unreliability of the parameters that are being determined.

In this manner, we have examined the effect of the local anisotropy and have come to the conclusion that it can be clarified by the rotational motion of the galaxy in the hypergalaxy if the rotational motion is placed on the isotropic expansion of our part of the Metagalaxy. Only research concerning extragalactic objects can answer the question of whether or not the last expansion is isotropic. On the other hand, the theory of an anisotropic and a non-homogeneous universe confirms (Ref. 45) that if the least anisotropy is now discovered, then it was significant in the past. In the light of this situation, the anisotropic and non-homogeneous Metagalaxy become more realistic in the ideal case of isotropy and homogeneity.

REFERENCES

1. Yu. P. Pskovskiy. Astronomicheskiy Zhurnal, 38, 521, 1961.
2. S. van den Bergh. Astrophysical Journal, 131, 215, 1960.
3. ______________ Astrophys. J. 131, 558, 1960.
4. ______________ Publications of Dunlap Observatory 2, No. 6, 1960.
5. Yu. P. Pskovskiy. Astron. Zh. 38, 1033, 1961.
6. K. Lundmark. Vistas in Astronomy, 2, 1607, 1956.
7. I. S. Shklovskiy. Astron. Zh. 37, 369, 1960.
8. Yu. P. Pskovskiy. Astron Zh. 38, 656, 1961.
9. H. Arp. Astrophys. J. 128, 166, 1958.
10. A. Sandage. Astrophys. J. 128, 150, 1958.
11. H. Arp, A. Sandage, C. Stephens. Astrophys. J. 130, 80, 1959.
12. H. Arp. Astrophys. J. 131, 322, 1960.
13. A. Sandage. Astrophys. J. 131, 610, 1960.
14. J. Fernie. Astrophys. J. 133, 64, 1961.
15. H. Arp. Astronomical Journal, 65, 404, 1960.
16. W. Baade. Mitteilungen Ast4onomische Gesellschaft, 1955. Hamburg, 1956, p. 51.
17. T. Schmidt. Zeitschrift fur Astrophysik, 41, 182, 1957.
18. A. Sandage. Astrophys. J. 127, 513, 1958.
19. N. Mayall, G. Kron. Astron. J. 65, 581, 1960.
20. G. Burbidge, M. Burbidge. Astrophys. J. 130, 629, 1959.
21. Yu. I. Yefremov. Astron. Zh. 26, 286, 1949.
22. H. Shapley, A. Ames. Harvard Annals, 88, No. 2, 1932.

REFERENCES (Cont'd)

23. M. Humason, N. Mayall, A. Sandage. Astron. J. 61, 97, 1956.

24. G. de Vaucouleurs. Astron. J. 58, 30, 1953.

25. ______ Vistas in Astronomy, 2, 1584, 1956.

26. ______ Astron. Zh. 36, 977, 1959.

27. ______ Astron. J. 63, 253, 1958.

28. K. F. Ogorodnikov. Dinamika Zvyezdnykh Sistem (Dynamics of Stellar Systems), Moscow, Gosudarstvennoye Fiziko-Matematicheskoye Izdatel'stvo, 1958.

29. Bolshaya Sovyetskaya Entsiklopediya (Great Soviet Encyclopedia), 2nd Edition, 27, p. 211, 1954.

30. E. Holmberg. Meddelanden Lunds Observatorie. II series, No. 136, 1957.

31. E. Pettit. Astrophys. J., 120, 418, 1954.

32. J. Stebbins, A. Whitford. Astrophys. J. 115, 284, 1952.

33. E. Holmberg. Annals Observatory at Lund, 6, 1937.

34. A. Reiz. Ann. Obs. Lund, 9, 1941.

35. J. Bigay. Annales d'Astrophysique, 14, 319, 1951.

36. P. P. Parenago. Astron. Zh., 22, 129, 1945.

37. Yu. P. Pskovskiy, A. S. Sharov. Reports of Gosudarstvennyy Astronomicheskiy Institut imeni P. K. Shternberga, No. 117, 1961.

38. S. van den Bergh. Publications of the Astronomical Society of the Pacific, 72, 312, 1960.

39. A. Hnatek. Z. F. Astrophys., 46, 224, 1958.

40. D. Mak Witti. Rasstoyanye i vremya v Kosmologii (Distance and Time in Cosmology). Symposium "Zvyezdnyye Sistemy", Moscow, Izdatel'stvo Inostrannoy Literatury.

41. K. Trumpler, H. Weaver. Statistical Astronomy. Berkeley and Los Angeles, University of California Press, 1953.

REFERENCES (Cont'd)

42. K. F. Ogorodnikov. Voprosy Kosmogonii, 1, 150, 1952.

43. A. L. Zelmanov. Doklady Akademii Nauk, SSSR, 102, No. 6, 815, 1956.

44. Yu. P. Pskovskiy. Astron. Zh., 37, 1056, 1960.

45. A. L. Zelmanov. Trudy VI Soveshchaniya po Voprosam Kosmogonii (Proceedings of VI Conference on Problems of Cosmogony), Izd-vo A. N. SSSR, 1959, p. 144.

THE INTERACTION OF THE GASEOUS COMPONENTS OF GALAXIES AND RADIO GALAXIES WITH AN INTERGALACTIC MEDIUM /44

By

N. S. Kardashev

Data on an intergalactic medium are briefly explained. New observation data on the deformation of galactic planes on a wavelength of 21 centimeters are given. Various interpretations of this effect are discussed. Obviously, this effect can be explained by an interaction with an intergalactic medium (hypothesis of Kahn and Woltjer). The problem on the interaction of extended radio galaxies with intergalactic matter are discussed. It is probable that such an interaction is observed in the case of NGC 5128.

Various considerations testify to the presence of an intergalactic diffuse medium. In the first place, most common are the concepts on the condensation process of galaxies which apparently continues at the present time [see, for example, (Ref. 1)]. In the second place, a number of cosmological hypotheses require a much greater mean density of matter in the universe than that which is obtained on the basis of the masses of galaxies. Thus, for example, in the equilibrial model of the universe and in the relativistic model by Einstein and de Sitter the mean density of matter = $3H^2/8\pi G$, where G is the gravitational constant. For the value of Hubble's constant $H = (4 \cdot 10^{17}\ \mathrm{sec})^{-1}$, we obtain $\rho \approx 10^{-29}\ \mathrm{g/cm^3}$ which is approximately two orders higher than the average density obtained throughout the galaxies. Finally, many authors notice a difference between the masses of groups and clusters of galaxies which are obtained according to the relationship of the mass to the luminosity (i.e., from their rotation) and according to the dispersion of velocities with the help of the virial theorem (Refs. 2-5). Even if the majorities of these objects are dynamically unstable or disintegrating (Refs. 1, 2), (which also testifies to their recent condensation) without a doubt there are many clusters of galaxies with a regular spherical form (of the type /45 of the known cluster in Coma Berenices), the stability of which it is difficult to doubt. It is extremely probable that in the last case, the stability is maintained by a diffuse medium of increased density exceeding in mass by once or twice the mass of the galaxies in the cluster.

Unfortunately, direct attempts to discover an intergalactic medium at the present time have not led to any reliable results. However, the experiments conducted indicate the upper limit of density and permit the voicing of quite likely concepts related to the physical properties of this substance.

In short, the results of various attempts to discover the intergalactic medium are reduced to the following. In accordance with the last work of Field (Ref. 6), the upper limit of density of the neutral intergalactic hydrogen is $nH < 2.3 \cdot 10^{-6}$ cm^{-3}, based on the supposition that the degree of ionization is not greater than 50 percent. This value was obtained from an investigation of the continuous radiation spectrum from the source Cygnus A in the region of the 21 cm line. The absorption by atoms of neutral intergalactic hydrogen should form a slight depression in the unbroken spectrum of the source, which flattens out due to the red shift into the wide absorption band which is limited by the insignificant radial velocity and the radial velocity of the Cygnus A source. Approximately the same value was reached in an attempt to discover the neutral hydrogen in the clusters of galaxies in Coma Berenices (Ref. 7), Corona Borealis, and Gemini (Ref. 8).

The direct discovery of the ionized component in the intergalactic medium is only slightly probable since the intensity of the fundamental processes of radiation of the strongly rarefied gas is proportional to the square of the electrons density in the presence of the expected temperature on the order of 10^6 degrees (Ref. 9), will be very small. The non-thermal radiation observed, such as that from the Metagalaxy on the whole (Ref. 10) as well as evidently that from certain clusters of galaxies (for example from the cluster in Coma Berenices) (Ref. 11) testify to the presence of an ionized gas. However, the non-thermal radiation only permits the determination of the magnitude of the magnetic field ($\sim 10^{-7}$ erg) and the density of the relativistic electrons.

The investigations concerning intergalactic absorption of light according to Zwicky (Ref. 12) evidently indicate a small quantity of dust within and in the environs of large clusters of galaxies. On the basis of the counting of the visible remote galaxies and clusters of galaxies through the cluster in Coma Berenices, the absorption was discovered in a photographic region about $0^m.7$ which for the diameter of the cluster ~ 14 Mps yields a dust density of $\sim 10^{-30}$ g/cm^3 (Ref. 5). If we assume the magnitude of the relation of gas to dust to be the same as in our galaxy (~ 200), then from the evaluation obtained we will find /46 the density of gas to be $\sim 2 \cdot 10^{-28}$ g/cm^3, which approximately increases two-fold the upper limit of density obtained from observations on a 21 cm wavelength for the same cluster.

In the same work by de Vaucouleurs (Ref. 5), there is information concerning the recent unsuccessful effort to discover an increased concentration of intergalactic stars in the environs of the nucleus of the cluster in Coma Berenices. Although, undoubtedly a small number of such objects exist (Ref. 12), it is uncertain whether they make up a significant part of the intergalactic medium.

The theoretical considerations indicate that if the intergalactic gas in early stages of evolution was ionized, then the probability of recombining is so small that it would remain ionized even up to the present time. All existing observations also evidently attest to the full ionization and high temperature of this medium. Further investigations concerning this are causing great difficulty.

However, within the past few years, it seems to us, an extremely favorable prospective was noted for research in the dynamic effects of the interaction of the intergalactic medium with the gas in the galaxy.

The work of Kahn and Woltjer (Ref. 9) was the first fundamental work in this direction. In this work, the possibility was shown of clarifying the deformation in the distribution of galactic interstellar hydrogen by an examination of the flux of intergalactic gas "blowing on" to the Galaxy.

THE DEFORMATION OF THE GASEOUS DISC OF THE GALAXY

As a result of the first experiments on the distribution of hydrogen in the galaxy, accomplished on the basis of the observations of radiation in the 21 centimeter line for both hemispheres, in Holland and Australia (Refs. 14, 15), it was revealed that the distribution of hydrogen in the galactic disc is not entirely planar. Between the galactocentric longitudes $L = 0 \div 120^{\circ}$ and at distances from the center of more than 10 kiloparsecs, the gas is distributed along the plane of the galaxy; on the opposite side it lies systematically lower than the plane. The general direction of the curve turned out to be close to the direction toward the Magellanic Cloud. It should be noted that all of these results were reached in the process of the very first investigations on the distribution of hydrogen in our galaxy. Because of these observations, it subsequently proved to be possible to define the position of the galactic plane. On the basis of this fact, a new system of galactic coordinates was introduced. The deformation of the gaseous disc was obtained as a secondary result. For a more detailed investigation of this phenomenon the Crimean station of the Physical Institute of the /47 Academy of Sciences of the USSR (Ref. 16) conducted observations with a radio telescope in the 21 cm wavelength on a specially coordinated program (Ref. 17) during 1960-61. On the part observed in the northern hemisphere, that of galactic plane ($\ell = 0 \div 240^{\circ}$), 13 points were found and differed in longitude by 20°. For each of the selected longitudes and $b = 0^{\circ}$*, the profile of the 21 cm line was written down. Then on

*We always use the new system of galactic coordinates (Ref. 18), and the galactic longitude is correspondingly counted from the direction of the center of the galaxy.

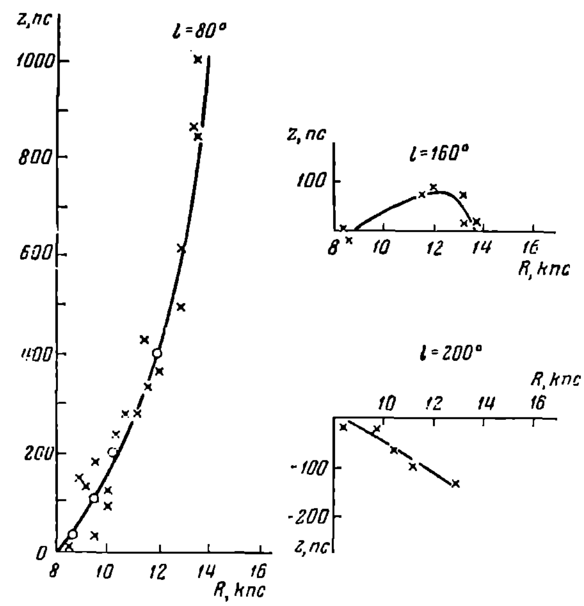


Fig. 1. Height of Inter-Stellar Gas Layers Above the Plane of the Galaxy in Relation to the Distance from the Center, for Longitudes of 80, 160, and 200°.

this profile, 6 to 8 fixed frequencies were selected, corresponding to the various distances from the Sun. On each of these frequencies, cross sections were made in the direction perpendicular to the plane of the galaxy, i.e., the antenna was shifted according to a precalculated program with an interval in most cases of 1 degree along the galactic latitude.

After the normal procedure of observation processing on the 21 cm wavelength (Ref. 19), the center of gravity of the distribution of intensity of hydrogen was found for each such cross section along the latitude, for each frequency, and for each of the selected longitudes. The entire program of observations was repeated three times. As a /48
result, for each of the longitudes, the dependence of the position of the center of gravity of the intensity of hydrogen layers on the distance from the center of the galaxy, was determined. In Fig. 1, the obtained dependencies for 3 longitudes are given. Along the axis of the abscissa, the distance R is projected from the center of the galaxy along the axis of the ordinates is projected the altitude z of the gaseous layers above the plane of the galaxy. Our data were shown by means of small crosses. For the longitudes $\ell = 80^\circ$, the analogous data reached by the Dutch investigators were shown by small circles. The comparison with the relief chart of the distribution of hydrogen in the galaxy obtained according to

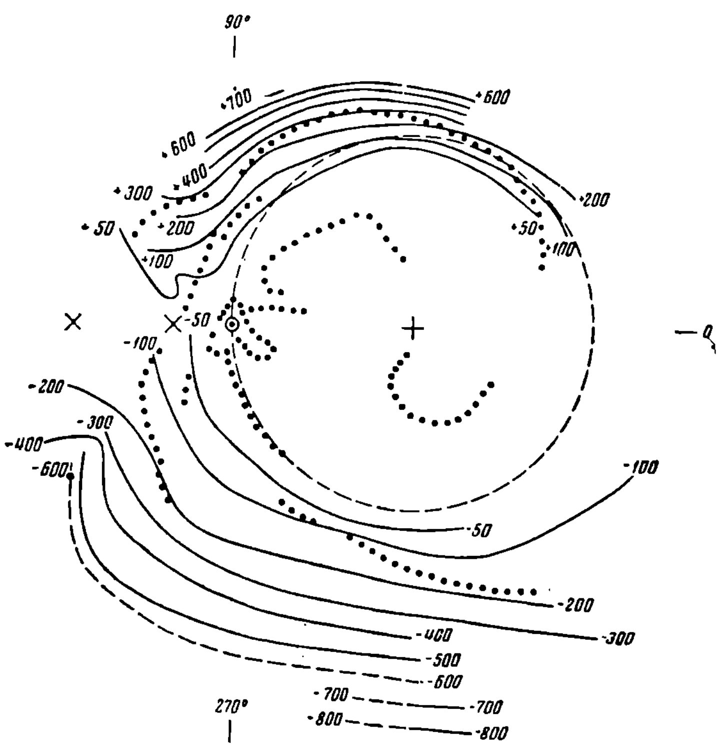


Fig. 2. Relief Chart of the Hydrogen Distribution in a Galactic Plane for the Law of Circular Rotation.

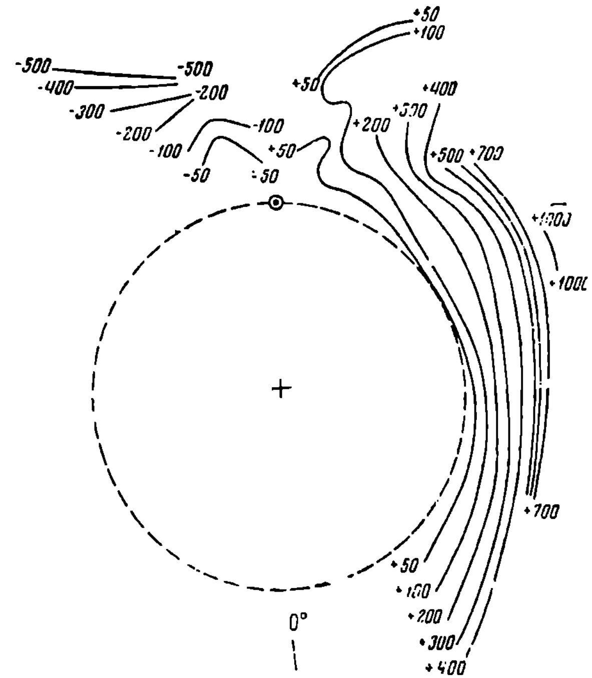


Fig. 3. Relief Chart of Hydrogen Distribution in a Galactic Plane for the Law of Rotation According to the Model by I. I. Genkin.

our data and the data reached concurrently by the Dutch and the Australians agrees well. Evidently, for individual longitudes, we have succeeded in tracing the effect being investigated to great distances from the center of the galaxy; the magnitude of the deformation approaches 1 kiloparsec for $\ell = 80^{\circ}$, and it has an unusually steep rise in distance. In Fig. 2 the relief chart of the distribution of hydrogen is shown with composition according to our observations. The altitudes are given in parsecs. This chart as in (Refs. 14, 15) was made on the assumption that the model of purely circular rotation of the galaxy is correct on the basis of the fact that all distances are also determined. However, recently a concept has appeared indicating that such an assumption is not completely valid. Actually, the purely circular motion from the point of view of stellar dynamics is only slightly probable, and evidently a majority of stellar systems, except for circular motion, should pulsate. I. I. Genkin (Ref. 20) made the charts of the distribution of hydrogen presupposing the existence of a radial velocity of the motion KR (R - the distance from the center of the galaxy). The best result was obtained for K = -2 km/sec which coincides well with the K-effect observed on distant stars. With this assumption, the clearly expressed spiral structure of the galaxy was obtained, contrary to results reached in the assumption of purely circular motion. We also made a relief chart for the model of I. I. Genkin making use of our observations and the observations of the Australian investigators.

By means of points are shown in Fig. 3 the spiral arms, which were obtained by I. I. Genkin according to interstellar hydrogen, and the continuous lines are iso high altitude curves. The numbers signify the height in parsecs. The question concerning the fact whether the hydrogen is compressed in relation to the plane of the galaxy or expanded cannot yet be considered conclusively solved, since the K-effect deals only with the gradient of radial velocity and not with the magnitude itself.

The results of the investigation of a 21 cm line in the direction of the galactic center deals with the affluence of hydrogen (Ref. 21). The presence of a line of absorption in the profile of the 21 cm line with positive radial velocity in the direction to the center of the galaxy (Ref. 22) possibly attests to the recession of the entire /50 system of gas and stars surrounding the Sun with a velocity of approximately 7 km/sec from the center.* Evidently, the Australian investigators (Ref. 23) arrived at analogous deductions. It is definite that the law of rotation essentially influences the magnitude of the effect of deformation in the distribution of gas; however, it is doubtful whether this gas will influence those rough evaluations which can be made at the present time concerning the interpretation of this effect.

---

*The profile of the 21 cm line in the direction to the anti center of the galaxy (Ref. 16) testifies to the presence of radial motion.

In spite of the fact that the tendency of deformation of the galactic disc coincides with the direction to the Magellanic Cloud, the estimates of the interaction of the influx gives the effect as being one to two times less (Ref. 24). Recently, the applied magnitudes of the distance and masses obtained for the Magellanic Cloud have changed somewhat. It has been agreed (Ref. 25) that the mass of the Major Magellanic Cloud is $(2.5 \pm 0.6) \cdot 10^{10}$ $M_{\odot}$ and that the distance is 63 kiloparsecs. Using this magnitude and the component of gravitational acceleration perpendicular to the galactic plane K_z (Ref. 26), we have determined the possible variation under the influence of the tidal disturbance, which together with the observed variation for $\ell = 80^{\circ}$ are shown in the table. The magnitude K_z is expressed in units $0.324 \cdot 10^{-9}$ cm/sec^2. The data obtained previously shows that the tidal forces may explain only an insignificant part of the observed effect. As it follows from the obtained relief charts, the fact that the deformation gradient is many times larger in a direction opposite to the Major Magellan Cloud than in the direction toward that cloud, testifies against the tidal interaction.

R, kiloparsecs	z_{tidal} parsecs	z_{obs}, parsecs	$K_{z\,observ.}$
8.2	0.5	0	0
9.2	0.8	80	2.6
10.2	1.5	110	2.2
11.2	3.2	250	2.3
12.3	6	450	2.0
14.3	21	1000	1.6
16.4	100	-	-

At the same time, in (Ref. 9) it was persuasively shown that the internal forces can hardly support the observed deformation. The differential rotation of the galaxy and oscillations near the plane with a frequency which is dependent on the distance to the center, should have completely annihilated such a deformation if it were formed in the moment of the condensation of the galaxy. It is irrefutable that the external /51 forces support the deformation. In connection with this, the Kahn and Woltjer assumption that the galaxy is located in a flux of intergalactic gas which is related to the local system of galaxies is very probable (or the hypothesis of Gold to which Kahn and Woltjer refer concerning the fact that

this gas is the intergalactic medium not related to the local cluster). In both cases, the approaching flux (Fig. 4) while traveling at velocities

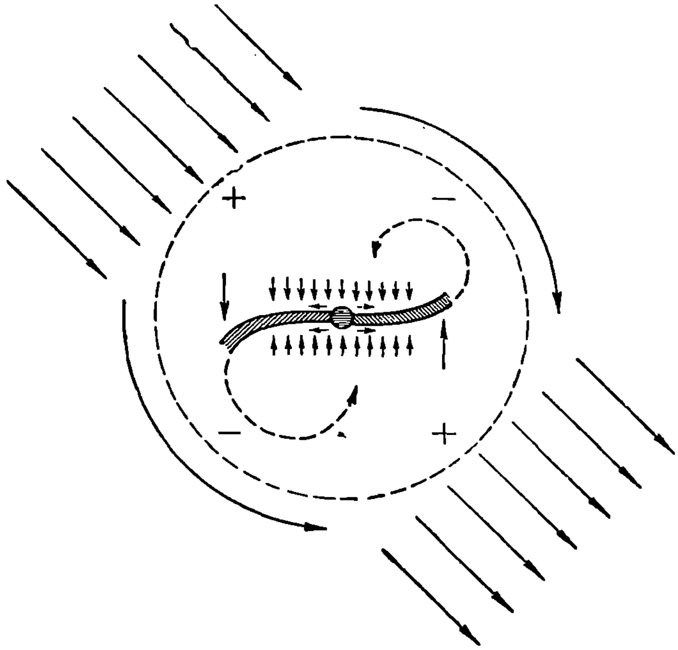


Fig. 4. Diagram of Interaction Between an Intergalactic Flux with a Galaxy.

approaching that of sound (with a density of the intergalactic medium 10^{-4} g/cm^{-3}, temperature $T = 0.5 \cdot 10^{6°}$, the speed of sound $a = 120$ km/sec) inflicts a positive pressure on the head and tail sides of the halo of the galaxy and a negative pressure in the perpendicular directions according to the Bernoulli law. The calculation which they have conducted shows that in the presence of the velocity of the flux of 60 km/sec, the arising difference in pressures is $\Delta p = 10^{-14}$ dynes $\cdot$ cm^{-2}. These pressures are transmitted to the gaseous disc and cause its deformation. From the correlation $\Delta p = -K_z \ell \rho$, where ℓ is the thickness of the gaseous disc ($\sim$ 200 parsecs), and ρ is its density, one can determine the magnitude of the deviation z; for R = 12 parsecs this deviation is approximately /52
100 parsecs, and approximately 300 parsecs for R = 14 kiloparsecs. The calculation of the resonance effect on the oscillation about the galactic plane and the rotation of the galaxy evidently permits to raise these

magnitudes to 150 parsecs at a distance of 12 kiloparsecs and 900 parsecs at a distance of 14 kiloparsecs. This is found to be in agreement with observations.

There are certainly doubts regarding the application of the equation of Bernoulli to the problem of the circumference of the galactic halo. It is uncertain whether the halo has such a regular spherical form; according to the isophotes of radio emission of M 31 (Ref. 11) it is clear that this is far from such. At the same time there is a basis to consider that the magnetic field of the galaxy extends far out in the intergalactic medium (Ref. 27). All of this considerably complicates the theoretical analysis. However, the calculations of Kahn and Woltjer apparently reflect the fact that the energies of such a type of interaction are sufficient for explaining the observed deformation and this, it appears to us, is the basic point. It seems quite probable to us that the indicated interaction can also guarantee the ejection of the gas from the periphery of the galactic disc in the halo. As is known, the hydrogen escape of $\sim M_{\odot}$ one solar mass per year is observed from the center of the galaxy along the plane (Ref. 21). Recently it has been discovered (Ref. 28) that from the side of the galactic folds in the surroundings of the Sun the hydrogen flows down to the plane of the galaxy. The magnitude of the flux is $\sim 10^{-7}$ $M_{\odot}$ per year per square parsec of the plane of the disc with the velocity of 12 km/sec. It is possible to show using the energy values that the continual motion of the galaxy in the flux of the intergalactic medium capably guarantees such circulation. It would be interesting to examine the problem concerning the interaction of galaxies with the medium in the presence of supersonic velocities. The energies of such processes are so great that possibly it can secure even the process of the ejection of gas from the galaxies as is the case in radio galaxies or even the crushing of galaxies. In connection with this it is possible that the formulation of the Magellanic Clouds is the result of the same cause.

At the present time, certain other galaxies are also known which apparently possess deformation of the gaseous disc. The Burbidges (Ref. 29) (see Fig. 5) noted that in NGC 5866 a spiral galaxy observed almost exactly from the edge a thin strip of dust matter with a thickness of 2-3" is slightly inclined at an angle of 2° relative to the major axis which is being determined by the distribution of the luminosity. This effect in our galaxy is approximately twice as great. Here it is again visible that the deformation has a non-effluvial character since the stellar component of the galaxy does not participate in the inclination. It seems to us that the investigation of similar objects can give much information concerning the velocities of motion of the intergalactic medium in various points of the universe.

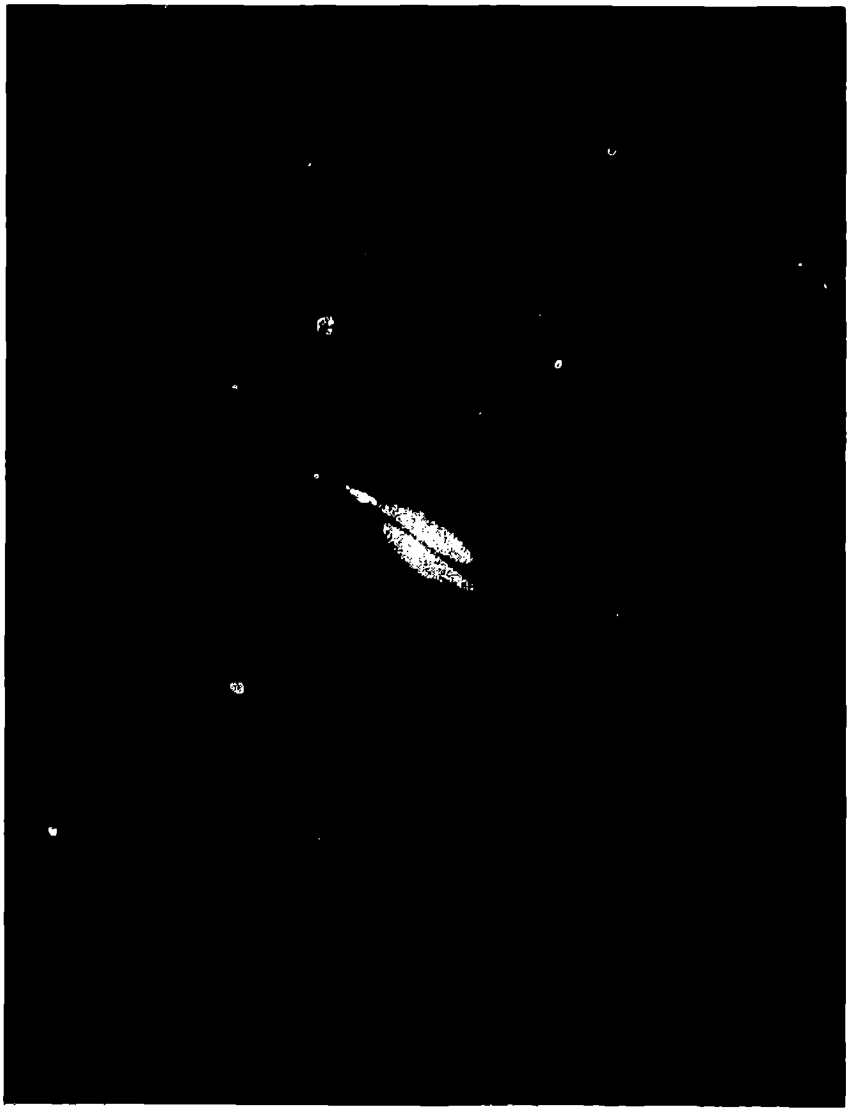


Fig. 5. Deformation of a Dust Belt in the Galaxy NGC 5866 (SO). Distance 12 Mparsecs, Thickness of Dust Belt 120-180 parsecs, Inclination $\sim 2^{\circ}$ (H = 75 km/sec · Mparsecs).

THE INTERACTION OF THE INTERGALACTIC MEDIUM WITH RADIO GALAXIES /54

Thanks to the successes of radio astronomy at the present time, it is known that a majority of the sources of non-thermal cosmic radio emission are connected with extremely remote galaxies. These objects which have received the name radio galaxies for their unusually powerful radio emission can be subdivided into two types. The dimensions of the radio emitting regions of certain of these are found within the optically observed galaxy; for example, such a galaxy is NGC 1068. Others and the most intense have regions of radio emission situated far outside the limits of the optically observed object. A typical example of such a type is the source Cygnus A. There are quite a few objects of the mixed type: possessing both a source with small angular dimensions and a source of greater extent. A characteristic example of such a type is the galaxy

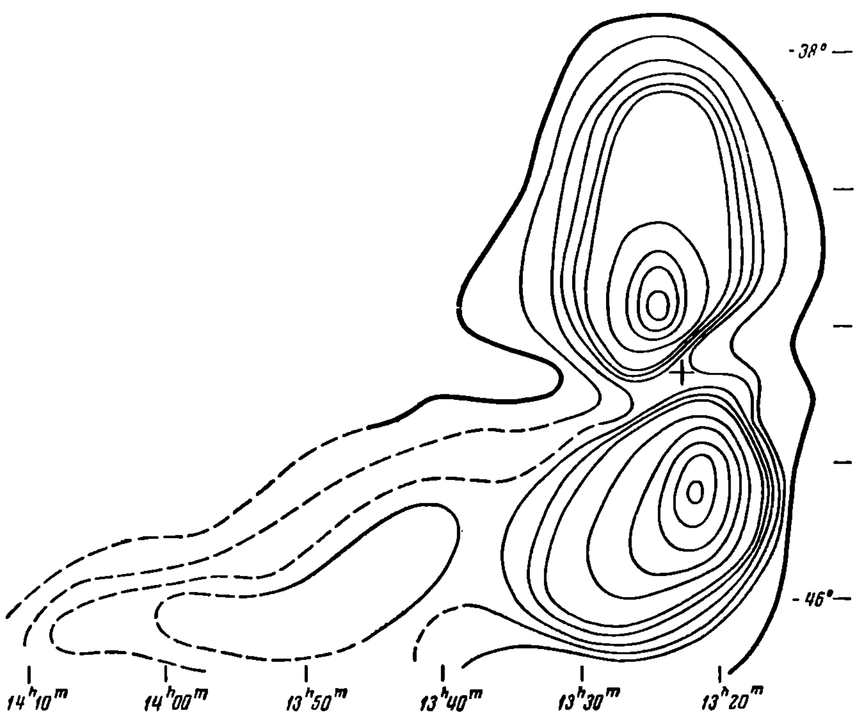


Fig. 6. Radio Emission Isophotes on a Wavelength of 21 cm.

NGC 5128. The regions of radio emission of this source extend to a distance of 1 megaparsec and possibly even farther. Such formations possess a completely insignificant mass in the presence of giant dimensions, and therefore the effects of the dynamic interaction with the intergalactic medium is for them especially great. In Fig. 6 are shown the isophotes of the radio emission of the expanded source NGC 5128 obtained by Bolton and Clark (Ref. 30). A small circle in the center shows the optically observed object. Accepting the distance to be 5 megaparsecs, we will find the volume of the expanded source $2 \cdot 10^{71}$ cm^3 for each cloud, and at a density 10^{-29} g $\cdot$ cm^{-3} its mass will be 10^9 $M_\odot$. At the same time, when a cross section of $4 \cdot 10^{47}$ cm^2 the collective mass of the flux which "blows out" with a velocity of 10^3 km/sec for 200 years will have the same magnitude. Therefore, it is extremely probable that the asymmetry visible in Fig. 6 in the location of the expanded source is connected with the effect of its deceleration in the galactic medium. The character itself of the isophotes also apparently deals with the presence of the flux of gas moving from the side of the least /55

Photograph of Radio Galaxy NGC 5128 .

direct ascension. On the photograph of the nebula the "tail" extended/56 along the direction of the expansion of the dust matter attracts attention. This tail is drawn (Fig. 7) close to the intersection with the isophotes of radio emission of the plane of the galaxy. It is mentioned in (Ref. 31), that this detail lays along the same small circle on the sphere, as the known projection of radio emission isophotes, which departs from the plane of the galaxy.

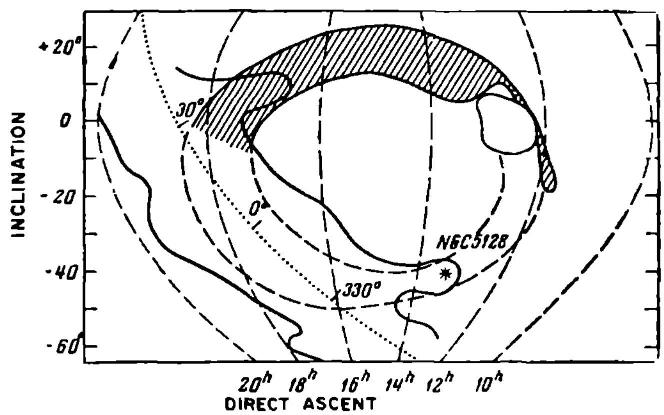


Fig. 7. Diagram of Radio Emission Isophotes on a Frequency of 600 mc Which Illustrates the Junction of Certain Details with the Radio Galaxy NGC 5128.

The interpretation of this formation as a cloud of a flared up super nova very close to the Sun, encounters a number of difficulties. First of all, it is necessary to mention the absence of any optical luminescence from this formation.

In connection with all that was stated previously, it is possible for us to suppose that all these details represent a small cloud of gigantic expansion (with the diameter of about 5 megaparsecs) ejected from the same source NGC 5128, and then shifted under the influence of the flow of intergalactic gas. The farthest optical and radio astronomical investigations of similar objects are extremely possible for the explanation of the nature of the intergalactic medium.

REFERENCES

1. V. A. Ambartsumyan. Trudy VI Sovyeshchaniya po Voprosam Kosmogonii (Proceedings of VI Conference on Problems of Cosmogony), Izdatel'stvo Akademii Nauk, SSSR, 1959, p. 12.

2. ______ Solvay Conference, Brussels, 1958, p. 241.

3. G. Burbidge, M. Burbidge. Astrophysical Journal, 130, 15, 1959.

4. G. de Vaucouleurs. Astrophys. J. 130, 718, 1959.

5. ______ Astrophys. J. 131, 585, 1960.

6. G. Field. Astronomical Journal 65, 547, 1960.

7. C. Muller. Bulletin of the Astronomical Institutes of the Netherlands 14, 339, 1959.

8. N. S. Kardashev. Astronomicheskiy Tsirkular 1961.

9. F. Kahn, L. Woltjer. Astrophys. J. 130, 705, 1959.

10. K. A. Shein. Radioastronomya, Paris Symposium 1958, p. 322. M. 1961.

11. Monthly Notices of the Royal Astronomical Society, 119, 347, 1959; C. Seeger, G. Westerhout, R. Conway. Astrophys. J. 126, 585, 1957; M. Large, D. Mathewson, C. Haselan. Nature 183, 1663, 1960.

12. F. Zwicky. Morphological Astronomy, Heidelberg 1957.

13. G. Field. Astrphys. J. 129, 536, 1959.

14. G. Westerhout, Bulletin of the Astronomical Institutes of the Netherlands 13, No. 201, 1957.

15. J. Öort, S. Kerr, G. Westerhout. Monthly Notices Roy. Astron. Soc., 118, 379, 1958.

16. N. S. Kardashev, T. A. Lozinskaya. Astronomicheskiy Zhurnal 39, No. 1, 1962.

17. R. L. Sorochenko, B. M. Chikhachev. Proceedings of Vth Conference on Problems of Cosmogony, Izd-vo AN SSSR, 1956, p. 546.

18. A. Blaauw, C. S. Gum, J. L. Pawsey, G. Westerhout. Monthly Notices Roy. Astron. Soc. 121, 123, 1961.

19. C. Muller, G. Westerhout. Bull. Astron. Inst. Neth. 13, No. 475, 151, 1957.

REFERENCES (Cont'd)

20. I. I. Genkin. Astron. Zh. 38, 000, 1961.

21. G. Rougoor, J. Öort. Proceedings of National Academy of Sciences of the U.S.A., 46, 1, 1960.

22. S. A. Muller. Radioastronomye, Paris Symposium 1958, p. 359, M. 1961.

23. F. Kerr. Astron. J. 66, 37, 1961.

24. F. Burke. Astron. J. 62, 9, 1957; F. Kerr, J. Haindman, M. S. Carpenter. Nature 180, 677, 1957.

25. G. de Vaucouleurs. Astrophys. J. 131, 265, 1960.

26. E. Hill. Bull. Astron. Inst. Neth. 15, No. 494, 1, 1960.

27. F. Hoyle, J. Ireland. Monthly Notices of the Royal Astronomical Society 120, 173, 1960.

28. H. Hindman, McGee, Carter, F. Kerr. Astron. J. 66, 37, 1961.

29. E. Burbidge, G. Gurbidge. Astrophys. J. 131, 224, 1960.

30. J. Bolton, B. Clark. Publications of the Astronomical Society of the Pacific 72, 29, 1960.

31. R. Brown, R. Davies, C. Hazard. Observatory 80, 191, 1960.

ON THE GRAVITATIONAL PARADOX

/58

By

A. Ya. Kipper

A gravitational paradox occurs when the laws of gravity are extended to an infinite universe. The law of gravitation is expressed by differential equations for the metric tensor $g_{\mu\nu}$ of the general theory of relativity or by a differential equation for the potential Ω in a non-relativistic approximation. In this type of formulation the law of gravitation is a local law; i.e., it is given for an infinitesimal time-space volume.

In order to obtain a law within the boundaries for the infinite universe as a whole from the local law of gravitation for finite regions, it is necessary to integrate the field equations.

The procedure of integration includes the definition of boundary conditions. In the non-relativistic theory the field equations coincide with the Poisson equation, and the boundary conditions, which are essential for its solution, present the requirement that at infinite remote points in space the density of matter should approach zero. But the actual universe is not constructed in such a manner that would permit the boundary conditions to be fulfilled, and we encounter a contradiction known as the gravitational paradox. Consequently, the gravitational paradox is a result obtained in process of cosmological problems.

An attempt is made in this paper to solve the Poisson equation without establishment of boundary conditions. The Poisson equation

$$\Delta\,\Omega = -\,4\pi\varkappa\rho \tag{1}$$

is usually solved by means of the integral

$$\Omega = \lim_{V\to\infty}\left\{\varkappa\int_{(V)}\rho\,(\zeta)\,\frac{d^3\zeta}{|x-\zeta|}\right\} \equiv \varkappa\int_{(\infty)}\rho\,(\zeta)\,\frac{d^3\zeta}{|x-\zeta|}\,. \tag{2}$$

If such a limit exists, the integral actually presents a solution to the equation. However, in order for this limit to exist, it is

necessary that the density of matter ρ (ζ) diminish quite rapidly at remote points in space. If these boundary conditions are not fulfilled, integral (2) diverges and the gravitational paradox occurs.

The divergence of integral (2) signifies that it does not have a numerical value. However, for certain problems among them and problems of a cosmological nature the divergent integrals can be utilized. In this article the divergent integrals are used as a result of the generalization in the concept of integrals which have an infinitely extensive integration. The generalized integral is

/59

$$\int_{(\upsilon_\infty)} \rho\ (\zeta)\ \frac{d^3\zeta}{|x - \zeta|} \qquad (3)$$

and is known as an integral having an actually infinite extent of integration. By this description it is stressed that the procedure of calculating the limit according to formula (2) is not included in the determination of the integral.

A generalized integral is not a mathematical object which can be expressed in actual numbers. However, generalized integrals do yield a corresponding analysis and prove to be convenient for a presentation of the potential Ω of universal gravitation. Generalized integrals solve the problem of the gravitational paradox.

A physical interpretation of the results obtained in the first part of this article is given in the second part. It has been shown that two systems of measuring space and time should exist. These systems are called gravitational and atomic. It should be noted that the existence of such systems has been postulated earlier by Milne. Not long ago Blakesley developed a hypothesis for two varieties of time.

INTRODUCTION

/60

For infinity as a whole, the gravitational paradox becomes a contradiction in the case of the application of the laws of gravitation of finite regions of the universe. The law of gravitation is expressed by the differential equations for the metric tensor $g_{\mu\nu}$ in the general theory of relativity, or it is expressed by the differential equation for the potential Ω in the non-relativistic approach. The presence of matter which causes the field is taken into consideration in the equations by means of the tensor of matter or the density of the material in a non-relativistic approach. The field equations combine the derivations of

the metric tensor $g_{\mu\nu}$ (or the potential Ω) along the coordinates at a given point with the tensor of the matter (or the density of the material) at this point. In such a formulation, the law of gravity appears to be a local law; that is, it is given for an infinitely small space-time volume.

In order to extend the law, in its most useful form, to the infinite universe as a whole from the local law of gravitation for finite regions of the universe, integration of the field equations is necessary. Here, difficulties arise, one of which is known as the gravitational paradox. The equations of the field themselves are free from a similar type of contradiction and can be correctly recognized as reflecting the objective truth. The equations of the gravitational field in a non-relativistic approach give the equation of Poisson (0.1); that is, the equation is an elliptical type. For its simple solution, it is necessary to set the value of the potential Ω at a certain surface σ surrounding a certain volume V, or as it is often stated, it is necessary to set /61 definite boundary conditions on the surface σ. In order to arrive at a solution of the Poisson equation for the entire infinite space (that is, as $V \to \infty$), it is necessary to set the value of the potential Ω at an infinitely remote surface σ.

$$\lim_{\sigma\to\infty} \Omega = \text{const.}$$

On the basis of the Poisson equation (0.1) this is equal to the requirement

$$\lim_{\sigma\to\infty} \rho = 0.$$

This expresses the following condition: in the infinite universe, it is possible to assign a cirdumference to a certain sphere with a sufficiently large radius so that, at points outside the sphere, the density of matter will prove to be less than any small number assigned beforehand. The expression "boundary conditions on an infinitely remote surface σ" is not exact since the motion of an infinitely remote surface is not connected with a definite surface. It is appropriate to name the boundary conditions as $\sigma \to \infty$ the extreme conditions, since the corresponding extreme value of ρ (or Ω) is assigned by means of these conditions.

There is a definite solution to the Poisson equation if the density ρ satisfies the boundary conditions and, on the other hand, a solution is not available (save for certain exceptions; see, for example, the footnote

on page 67) if the boundary conditions are not satisfied. On the other hand, the universe, constructed so that some kind of boundary conditions for ρ might be satisfied, can be called infinite only conditionally. In the real universe, the density of matter ρ on the surface of $\sigma \to \infty$, does not approach zero. This contradiction is expressed in the fact that for the really essential infinite universe, the Poisson equation has no solution. This contradiction is known as the gravitational paradox. According to what was stated previously, the gravitational paradox arises as a result of the limitation of the mathematical apparatus still being used for the solution of cosmological problems.

A. Einstein has also pointed out the incompatibility between boundary conditions and the physical essence of the universe. It is possible to solve field equations within the general relativity theory, without determining boundary conditions, assuming that the universe is space limited. However, in spite of the logical sequence and beauty of the models of the enclosed universe, from the point of view of philisophical knowledge they seem unsatisfactory. In the present investigation, an attempt is being made to set up a hypothesis of the infinity of the universe which was partially established by brilliant achievements /62
in the general theory of relativity, but not without doubt. An attempt is being made to solve the equation of the gravitational field in non-relativistic approach without the boundary condition: $\rho \to 0$ as $\sigma \to \infty$. Basically, the concept of a solution to the problem is contained in the following.

The usual method of solution of the Poisson equation

$$\triangle \Omega = - 4\pi \varkappa \rho \tag{0.1}$$

for all universal space is contained in the two steps of calculation. First of all it is assumed that the density of matter $\rho(x)$ is a function of the coordinates x_j of the points of space that are different from zero only in a certain finite volume V. Then

$$\Omega_V = \varkappa \int_{(V)} \rho \; (x') \frac{d^3x'}{|x - x'|} \tag{0.2}$$

is the solution of equation (0.1). It is further considered that the infinite universe is something like a copy of finite systems of gravitating bodies. The only difference between them is the fact that volume V of the entire universal space is infinite. This can be summarized as, that for the infinite universe,

$$\Omega = \lim_{V\to\infty} \left\{\varkappa \int_{(V)} \rho\ (x') \frac{d^3x'}{|x - x'|}\right\} \equiv \varkappa \int_{(\infty)} \rho\ (x') \frac{d^3x'}{|x - x'|}\ , \qquad (0.3)$$

whereby the integral in the last expression is understood to be the limit of the endlessly increasing finite volumes of integration V. If such a limit exists, then the integral (0.3) will be actually the solution of the Poisson equation (0.1).

The convergence of the integral (0.3), or the existence of a corresponding boundary, requires definite conditions for the subintegral expression; that is, for the density function $\rho(x')$, which is reduced to the requirement that $\rho(x')$ decrease sufficiently with distance in the infinitely remote points. In other words, $\rho(x')$ as a function of the coordinates, should satisfy the corresponding specified boundary conditions.* If the specific conditions are not fulfilled, the integral (0.3) diverges and we have a case of the gravitational paradox.

The divergence of the integral (0.3) signifies that this integral does not have a definite numerical value. However, the divergent integrals can be utilized for the solution of certain cosmological problems among them. If it is expedient to determine the mathematical /63 meaning of the symbols which are being called generalized integrals, then,independent of the convergence of divergence of the integrals, it is possible to express the physical laws of connections by these symbols. The generalized integral is as follows:

$$\int_{(V_\infty)} \rho\ (x') \frac{d^3x'}{|x - x'|} \qquad (0.4)$$

and we shall also define it as an integral with an actually infinite volume of integration. This expression emphasizes the fact that the procedure for calculating the limit for an endlessly increasing finite volume of integration $V \to \infty$ is not included in the definition of the integral. Since the volume of universal space corresponds to the magnitude of integration, the determination of the integral (0.4) signifies that the infinity of the universe is examined as actual infinity.

*The fulfillment of boundary conditions when $\rho(x')$ is satisfactory, but not essential for the convergence of the integral (0.3). There are available functions of $\rho(x')$ which do not satisfy the boundary conditions, in the presence of which the integral (0.3) converges. A known example of such a $\rho(x')$ is the matter density functions by Charles in his model of the universe. However, the Charles hypothesis is of little value.

Two concepts of infinity are sometimes distinguished in the theory of numbers. These concepts are known as potential and actual infinities, whereby extremely deep meaning is attached to these terms. Infinitive numbers are considered to be actually known, if the definition of numbers is a criterion, by means of which it is possible to determine definitely whether a specified object (really existing or abstractly conceived) is included in these numbers. Thus, for example, the evaluation of "universal space" is numbers of its points -- a determination which is actual since it is considered that a point in space is a completely determined object. Actually, determined endless numbers are known to be actually endless.

On the other hand, the infinity of universal space can be understood as potential infinity. If a certain suitably large volume of space V is set up, there will always be a potential possibility to set up a volume V' completely including volume V. The concept of potential infinity of space is expressed by this potential possibility.

It is apparent that the image of the world as an actually infinitive problem of boundary conditions does not make any sense at all. The formulation of the specific conditions contains latently the presentations of the infinity of the universe in the sense of potential infinity. The mutual bond between the concepts of actual and potential infinities is still not completely solved by mathematical logic. However, it seems to us that the essence of the universe is encompassed in the concept of actual infinity much deeper than in the concept of potential infinity. With respect to this, the determination of the integral (0.4) as the integral with an actual infinite magnitude of integration V_∞ gives more /64 than the integral (0.3), which is connected with the calculation of the limit and with the presentation concerning potential infinity of the universal space.

The generalized integral (0.4), the exact determination of which is presented in the following paragraphs, is not a mathematical object expressed in real numbers. However, the generalized integrals are given over to the corresponding analysis and are no less suitable for the presentation of the potential Ω of universal gravitation and usual functions of the coordinates of the points of space.

The basic idea of the solution of cosmological problems with the help of mathematical formalism distinct from the formalism of classical physics is, generally speaking, not new. Quantum mechanics applies a mathematics which is qualitatively distinct from the mathematics of traditional physics. Since cosmic dimensions differ so significantly from everyday dimensions, such as do the hydrophysical dimensions, there

is nothing astonishing about the fact that in cosmology, i.e., in physics of super-large dimensions, we encounter circumstances compelling the application of calculative methods which are distinct from the usual methods.

The non-relativistic theory of gravitation is based on the field equation (the Poisson equation) and on the equation of the motion of matter in the field. In this investigation are generalized also the equations of the motion of particles in that field, in addition to Ω potential of universal gravitation. Instead of common values x_j it is expedient to utilize in the capacity of coordinates, the generalized integrals of the following form

$$\xi_j = \frac{3}{4\pi} \int_{(V_\infty)} a\ (t) \cdot \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta, \qquad (0.5)$$

whereby the expression (0.5) is also a reorientation formula, which combines the system of coordinates (x_j) expressed in common values, with the system of coordinates (ξ_j) expressed in generalized integrals. In the logical theory, it is also necessary to apply two systems of time t and τ so that (x, t) and (ξ, τ) are qualitatively diverse systems of coordinates of space-time. In the non-relativistic theory, t and τ are the usual quantities.

The second half of the present investigation is devoted to the physical interpretation of results reached in the first half of the work. An attempt is being made to prove that the quantitatively diverse systems of coordinates (x, t, and ξ, τ) mentioned earlier reflect the actual existing distinction. Of course, it is not necessary to believe that there are two spaces and two times somehow pushed into one another. The space-time continuum is one, but this continuum can be measured by /65 instruments of diverse physical nature. For example, two types of clocks exist (called atom and gravitational clocks) and those and others are valid, but in view of the difference of their physical nature, they cannot be synchronized.

The existence of two systems for measurement of time and space in this investigation, is based also on the Weyl principle, which asserts that it is not expedient to postulate the original existence of devices for the measurement of distance and time lapse, in order to construct a general relativity theory. Such a device may be constructed on the basis of an existing theory. It is also necessary to note that in his

time, Milne postulated the existence of two systems for the measurement of time. Recently a two-time hypothesis was developed by Blakesley.

The atom and gravitational clocks are valid; therefore, there is no physical reason to consider some of them to be fundamental. But just the same, if you consider, for example, the atom time to be valid then you will come to the conclusion that the constant of gravitation will be changed. This is the well known Dirac hypothesis. In the present research, it will be regarded as a consequence of the theory of universal gravitation in a non-relativistic approximation. Another consequence are the results following, which in the atom system of measuring space and time, a vacuum should have the characteristics of outstanding gravitating material. In conclusion, several more deductions having the ability to possess a cosmogonical value are being examined.

* * *

Following are some symbols and definitions which are used in this paper:

$\varkappa$, G - constant gravitations,

ρ - density of matter,

V - volume of space,

Ω - gravitational potential,

x_j, ξ_j - space coordinates,

t, τ - time,

$d^3x = dx_1\, dx_2\, dx_3$,

$|x| = (x_1^2 + x_2^2 + x_3^2)^{1/2}$,

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} .$$

Example of summation rule:

$$x_j x_j = \sum_{j=1}^{3} x_j x_j = |x|^2 .$$

If the coordinates are written without an index, then their entire complex is taken into consideration. For example, /66

$$f(x) = f(x_1, x_2, x_3).$$

§1. The Theory of Generalized Integrals

The symbols

$$J = \int_{(V_\infty)} \{\rho_1(\zeta) F_1(x, \zeta)$$

$$+ \rho_2(\zeta) F_2(x, \zeta) + \cdots + \rho_n(\zeta) F_n(x, \zeta)\} d^3\zeta, \quad (1.1)$$

are objects of the theory of generalized integrals. The mathematical meaning of these symbols is established by an axiom or a number of definitions. The continuous functions in the subintegral expression

$$F_k(x, \zeta) \equiv F_k(x_1, x_2, \cdots, x_m, \zeta_1, \zeta_2, \zeta_3), \quad (1.2)$$

which were defined for all values of variables in the integration ζ_1, ζ_2, ζ_3, where $-\infty$ to $+\infty$ are known as the nuclei of the integral. The functions

$$\rho_k(\zeta) \equiv \rho_k(\zeta_1, \zeta_2, \zeta_3), \quad (.13)$$

which should not necessarily be continuous are called density functions. The variables

$$x_1, x_2, \cdots, x_m$$

are known as the arguments of the integral.

The concept of the converging integral or of the integral in the usual meaning is considered to be known. The generalized integrals present a class of which a great number of all converging integrals is a subclass. The generalized integrals are of such a type that in their presence, the operation

$$\lim_{V\to\infty}\left\{\int_{(V)} \rho_1(\zeta)\, F_1(x, \zeta) + \right.$$

$$\left. + \rho_2(\zeta)\, F_2(x, \zeta) + \cdots + \rho_n(\zeta)\, F_n(x, \zeta)\right\} d^3\zeta$$

does not yield a well-defined result. With the determination of the generalized integrals, the operation of the transition to the limit $V \to \infty$ is not utilized; however, it serves as a method for the isolation of the converging integrals from the general class. The generalized integral, if it does not belong to the subclass of converging integrals, is not represented as a number. If we wish to utilize the generalized integrals in a physical theory, it is necessary to determine to which mathematical functions they are subordinate and also define the concepts of quality, identity, etc., of the generalized integrals.

The converging integral is first of all a definite sum. This /67 is expressed in the fact that the linear combination of two integrals

$$a_1 \int \rho_1(\zeta)\, F_1(x, \zeta)\, d^3\zeta + a_2 \int \rho_2(\zeta)\, F_2(x, \zeta)\, d^3\zeta \equiv$$

$$\equiv \int \{a_1\rho_1(\zeta)\, F_1(x, \zeta) + a_2\rho_2(\zeta)\, F_2(x, \zeta)\}\, d^3\zeta \qquad (1.4)$$

is an integral. Regardless of whether the integral will be converging or not, we shall require in all cases the fulfillment of condition (1.4), so that the generalized integrals have the property of a sum. Along with the fulfillment of (1.4) we shall also consider as correct the formula

$$a_1 \int_{(V_\infty)} \rho_1(\zeta)\, F_1(x, \zeta)\, d^3\zeta \equiv \int_{(V_\infty)} \{a_1\rho_1(\zeta)\, F_1(x, \zeta) +$$

$$+ a_2\rho_2(\zeta)\, F_2(x, \zeta)\}\, d^3\zeta - a_2 \int_{(V_\infty)} \rho_2(\zeta)\, F_2(x, \zeta)\, d^3\zeta\,, \qquad (1.5)$$

which we consider equivalent to formula (1.4) and which also expresses the possibility of transferring the integral from one part of identity to another.

If in formula (1.4) $a_1 = a_2 = 1$, then there is an addition; if $a_1 = - a_2 = 1$, then there is a subtraction; and if $a_1 \neq 0$, $a_2 = 0$, then there is a multiplication by the value a_1. The subtraction operation is defined by the "minus" sign in formula (1.5).

The difference of two diverging integrals can prove to be a converging integral. This will always occur when the nuclei and the density functions of both integrals are identical. As a result of subtraction, we obtain an integral with a subintegral expression, which is identical to zero. Such an integral converges and is identical to zero. Herefrom, it becomes possible to define two integrals as identical if their nuclei and functions of density are identical.

However, such a definition of identity is not sufficient. Therefore, we assume that two integrals are identical if their difference is a converging integral identical to zero. Since a converging integral can be identical to zero without the subintegral expression being identical to zero, this definition is more general than the original definition.

Two integrals with arguments of concrete values are equal if their difference is a converging integral equal to zero. Two dissimilar integrals can be equal if the values of the arguments are known. However, when equality is observed at all values of the arguments, then the integrals will be identical.

If there are two integrals J_1 and J_2, the difference of which is a converging integral with a value greater than zero (that is, a positive number), then /68

$$J_1 - J_2 > 0 \text{ or } J_1 > J_2, \tag{1.6}$$

i.e., we define J_1 to be greater than J_2. In a similar way, we define

$$J_2 - J_1 < 0 \text{ or } J_2 < J_1. \tag{1.7}$$

It is not possible to establish the correlations (1.6) or (1.7) for all integrals of the general class. A great number of the generalized integrals can not be completely regulated, but only partially.

A determination of a linear combination with integrals in accordance to formula (1.4) shows that the generalized integrals are mathematical objects known in the operators theory as lineals. Since a number of generalized integrals can be partially regulated, it is possible to consider the algebra of integrals as an algebra of semi-regulated linear spaces. This algebra has been well developed in mathematics.

We will not concern ourselves, in this paper, with the development of a mathematical theory of generalized integrals, although this theory is of some interest to cosmology. Only integrals of a specific type will be considered, and specifically

$$\int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{d^3\zeta}{|x - \zeta|}\ , \tag{1.8}$$

and also the integrals

$$\int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{\zeta_j - x_j}{|x - \zeta|^3}\ d^3\zeta \tag{1.9}$$

and still further, the integrals

$$\frac{3}{4\pi} \int\limits_{(V_\infty)} a\ (t)\ \frac{x_j - \zeta_j}{|x - \zeta|^3}\ d^3\zeta, \tag{1.10}$$

in which the density function a(t) is not dependent upon the variables of integration but might depend on a definite parameter t.

An example of the algebraic functions with integrals of the type (1.10) is the following variation

$$\frac{3}{4\pi} \int\limits_{(V_\infty)} a\ \frac{x_j - \zeta_j}{|x - \zeta|^3}\ d^3\zeta - \frac{3}{4\pi} \int\limits_{(V_\infty)} a\ \frac{x'_j - \zeta_j}{|x' - \zeta|^3}\ d^3\zeta \equiv$$

/69

$$\equiv \frac{3}{4\pi} \int\limits_{(V_\infty)} a\ \left\{\frac{x_j - \zeta_j}{|x - \zeta|^3} - \frac{x'_j - \zeta_j}{|x' - \zeta|^3}\right\}\ d^3\zeta = a\ (x_j - x'_j)\ , \tag{1.11}$$

which proves to be a converging integral.

Since two integrals are identical if their difference is a converging integral equal to zero, then we obtain

$$a' \int_{(V_\infty)} \frac{x'_j - \zeta_j}{|x' - \zeta|^3} d^3\zeta + a'' \int_{(V_\infty)} \frac{x''_j - \zeta_j}{|x'' - \zeta|^3} d^3\zeta \equiv$$

$$\equiv (a' + a'') \int_{(V_\infty)} \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta, \qquad (1.12)$$

where

$$x_j = \frac{a'x'_j + a''x''_j}{a' + a''},$$

whereby a' and a'' are constants and $a' + a'' \neq 0$.

The usual values can be considered as converging integrals. The following formula exists:

$$a \frac{3}{4\pi} \int_{(V_\infty)} \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta + b_j \equiv$$

$$\equiv a \frac{3}{4\pi} \int_{(V_\infty)} \frac{x_j + \frac{b_j}{a} - \zeta_j}{|x + \frac{b}{a} - \zeta|^3} d^3\zeta, \qquad (1.13)$$

where a and b_j are numbers.

Let us assume that in the integral

$$\int_{(V_\infty)} \rho\ (\zeta) \frac{d^3\zeta}{|x - \zeta|} \qquad (.114)$$

the density function does not satisfy the boundary conditions and that the integral does not converge. However, existence of the constant $\bar{\rho}$ is assumed to have a characteristic such that the integral

$$\int\limits_{(V_\infty)} (\rho - \bar{\rho}) \frac{d^3\zeta}{|x - \zeta|} \equiv \int\limits_{(\infty)} (\rho - \bar{\rho}) \frac{d^3\zeta}{|x - \zeta|} = \omega \; (x) \tag{1.15}$$

converges and determines definitely the function ω (x). With the convergence of integral (1.15) the following integral also converges

$$\int\limits_{(\infty)} (\rho - \bar{\rho}) \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta = F_j \; (x) = - \frac{\partial\omega}{\partial x_j}, \tag{1.16}$$

where F_j (x) is a definite function of the argument x_j. The existence /70 of the boundary

$$\lim_{V\to\infty} \left\{\frac{1}{V} \int\limits_{(V)} \rho \; (\zeta) \; d^3\zeta\right\} = \bar{\rho}. \tag{1.17}$$

appears to be a sufficient (but not essential) condition for the convergence of the integrals (1.16). If the density function $\rho(\zeta)$ is a density function of matter in universal space, then $\bar{\rho}$ is the mean density of matter in the infinite universe.

Utilizing formula (1.13), it is possible to prove that

$$\int\limits_{(V_\infty)} \rho \; (\zeta) \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta \equiv \bar{\rho} \int\limits_{(V_\infty)} \frac{x_j + \frac{1}{\bar{\rho}} F_j \; (x) - \zeta_j}{\left|x + \frac{1}{\bar{\rho}} F - \zeta\right|} d^3\zeta, \tag{1.18}$$

where F_j (x) is a definite function which is determined by formula (1.16).

By means of formulae (1.13) and (1.18) it is possible to demonstrate the accuracy of the following correlations:

$$\frac{3}{4\pi} \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{x_j + \Delta x_j - \zeta_j}{|x + \Delta x - \zeta|^3}\, d^3\zeta \equiv \frac{3}{4\pi} \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta - \Delta x_j \rho\ (x) \equiv$$

$$\equiv \frac{3}{4\pi} \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta + \Delta x_k \cdot \frac{3}{4\pi} \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{\partial}{\partial x_k} \left\{\frac{x_j - \zeta_j}{|x - \zeta|^3}\right\} d^3\zeta\ , \tag{1.19}$$

$$\frac{3}{4\pi} \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{d^3\zeta}{|x + \Delta x - \zeta|} \equiv \frac{3}{4\pi} \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{d^3\zeta}{|x - \zeta|} -$$

$$- \Delta x_j \frac{3}{4\pi} \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta - \frac{1}{2} |\Delta x|^2 \cdot \rho\ (x) \equiv$$

$$\equiv \frac{3}{4\pi} \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{d^3\zeta}{|x - \zeta|} + \Delta x_k \frac{3}{4\pi} \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{\partial}{\partial x_k} \left\{\frac{1}{|x - \zeta|}\right\} d^3\zeta +$$

$$+ \frac{1}{2} \Delta x_j \Delta x_k \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{d^2}{\partial x_j \partial x_k} \left\{\frac{1}{|x - \zeta|}\right\}\ , \tag{1.20}$$

whereby the existence of the quantity $\bar{\rho}$, which is determined by formula (1.17), is assumed.

The boundary

$$\lim_{\Delta x_j \to 0} \frac{1}{\Delta x_j} \left\{ \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{x_j + \Delta x_j - \zeta_j}{|x + \Delta x - \zeta|^3}\, d^3\zeta - \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta \right\}$$

will be defined as the derivative of the integral /71

$$\int_{(V_\infty)} \rho\ (\zeta)\ \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta$$

according to argument x_j, and will be designated by the symbol

$$\frac{\partial}{\partial x_j} \int_{(V_\infty)} \rho\ (\zeta)\ \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta \text{ (not to be summarized by j).}$$

On the basis of formula (1.19) it is possible to see that such a boundary exists and that $\frac{\partial}{\partial x_j} \int_{(V_\infty)} \rho\ (\zeta)\ \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta = - \frac{4\pi}{3}\, \rho\ (x)$ (not to be summarized by j).

The derivatives for the integral

$$\int_{(V_\infty)} \rho\ (\zeta)\ \frac{d^3\zeta}{|x - \zeta|}$$

are determined analogously.

It follows from formula (1.20) that:

$$\frac{\partial}{\partial x_j} \int_{(V_\infty)} \rho\ (\zeta)\ \frac{d^3\zeta}{|x - \zeta|} = \int_{(V_\infty)} \rho\ (\zeta)\ \frac{\zeta_j - x_j}{|\zeta - x|^3}\, d^3x, \tag{1.22}$$

$$\frac{\partial^2}{\partial x_j \partial x_k} \int_{(V_\infty)} \rho\ (\zeta)\ \frac{d^3\zeta}{|x - \zeta|} = - \frac{4\pi}{3}\, \delta_{jk} \cdot \rho\ (x), \tag{1.23}$$

where δ_{jk} is the Kronecker symbol. From formula (1.23) it follows that the integral

$$\Omega = \varkappa \int_{(V_\infty)} \rho \ (\zeta) \ \frac{d^3\zeta}{|x - \zeta|} \tag{1.24}$$

satisfies the Poisson equation

$$\Delta \, \Omega = - \, 4\pi\varkappa\rho.$$

Let us assume that in the integral

$$J_j \ (t) = \frac{3}{4\pi} \int_{(V_\infty)} a \ (t) \ \frac{x_j \ (t) - \zeta_j}{|x(t) - \zeta|^3} \, d^3\zeta \tag{1.26}$$

the density function a(t) does not depend upon the variables of the integration ζ_j but that it does depend upon the parameter t. Let us also assume that the arguments of integral x_j (t) also depend upon /72 t, so that the integral J_j (t) is some sort of function of t. We define

$$\frac{dJ_j}{dt} = \lim_{\Delta t \to 0} \frac{J_j \ (t + \Delta t) - J_j \ (t)}{\Delta t} \, .$$

By using the preceding formulae, it is possible to prove the validity of the following correlations:

$$\frac{dJ_j}{dt} = \frac{3}{4\pi} \int_{(V_\infty)} \frac{da}{dt} \, \frac{x_j - \zeta_j}{|x - \zeta|^3} \, d^3\zeta + a \, \frac{dx_j}{dt} \, , \tag{1.27}$$

$$\frac{d^2J_j}{dt^2} = \frac{3}{4\pi} \int_{(V_\infty)} \frac{d^2a}{dt^2} \, \frac{x_j - \zeta_j}{|x - \zeta|^3} \, d^3\zeta + 2 \, \frac{da}{dt} \cdot \frac{dx_j}{dt} + a \, \frac{d^2x_j}{dt^2} \, . \tag{1.28}$$

The integrals of type (1.10) will be used in the present investigation as coordinates of spatial points. We shall prove that this is possible.

Three dimensional space is a three dimensional continuum of its points. The reflection of this continuum on the continuum of real numbers is applied for the sake of convenient study of the continuum of the points. A set of three x_j, known as coordinates, is compared to each point if the numbers x_j, which correspond to infinitely close points differ only to an infinitesimal degree. In the case of a given space, it is possible to construct an infinitely great number of diverse systems of coordinates which are mutually bound by continuous conversions. The application of actual numbers of x_j in the capacity of space coordinates is not necessary. Any multitude of mathematical objects for which the definition of an infinitely small variation has been determined, is applicable for this purpose. The variation of the integrals

$$X_j = \frac{3}{4\pi} \int\limits_{(V_\infty)} \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta \text{ and } X_j' = \frac{3}{4\pi} \int\limits_{(V_\infty)} \frac{x_j' - \zeta_j}{|x' - \zeta|^3} d^3\zeta \tag{1.29}$$

according to formula (1.11) is a convergent integral and, consequently, a common quantity:

$$X_j - X_j' = x_j - x_j'. \tag{1.30}$$

We will arrange the quantities of integrals X_j in the following manner:

$$\left.\begin{aligned} X_j &< X_j' \text{ if } x_j < x_j'; \\ X_j &= X_j' \text{ if } x_j = x_j'; \\ X_j &> X_j' \text{ if } x_j > x_j'. \end{aligned}\right\} \tag{1.31}$$

Furthermore, let us assume that the quantities x_j are right angle /73 coordinates of points in Euclidean space. The distance of the points x_j and x_j' is

$$r = [(x_1 - x_1')^2 + (x_2 - x_2')^2 + (x_3 - x_3')^2]^{1/2} = |x - x'|. \tag{1.32}$$

On the basis of equation (1.30) we will have

$$r = |x - x'| = [(X_1 - X_1')^2 + (X_2 - X_2')^2 + (X_3 - X_3'^2]^{1/2} = |X - X'|. \quad (1.33)$$

and so the conformity

$$x_j \rightleftarrows X_j = \frac{3}{4\pi} \int_{(V_\infty)} \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta \quad (1.34)$$

is a mutually simple representation of the quantities x_j for the large number of integrals X_j. According to formula (1.31), the arrangement is preserved in the representation, and the knowledge of distance r is in agreement with formulae (1.32) and (1.33). Therefore, if the numbers x_j are, in essence, right angle coordinates of Euclidean space, then the generalized integrals X_j will also be coordinates of this space.

Let us assume that x_j and y_j are right angle coordinates of spatial points in actual numbers connected with the conversion formulae

$$x_j = \alpha_{jk}y_k, \quad y_k = \beta_{kj}x_j, \quad (1.35)$$

whereby

$$\alpha_{j\ell}\beta_{\ell k} = \delta_{jk}, \quad \beta_{j\ell}\alpha_{\ell k} = \delta_{jk}, \quad \alpha_{j\ell}\alpha_{\ell k} = \delta_{jk}, \quad \beta_{j\ell}\beta_{\ell k} = \delta_{jk}, \quad (1.36)$$

that is, the rotation of the coordinate axis takes place without displacing the origin of the coordinates. Furthermore, let us establish the conformity

$$x_j \rightleftarrows X_j = \frac{3}{4\pi} \int_{(V_\infty)} \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta, \quad y_j \rightleftarrows Y_j = \frac{3}{4\pi} \int_{(V_\infty)} \frac{y_j - \zeta_j}{|y - \zeta|^3} d^3\zeta. \quad (1.37)$$

We obtain

$$X_j = \frac{3}{4\pi} \int_{(V_\infty)} \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta = x_j + \frac{3}{4\pi} \int_{(V_\infty)} \frac{\zeta_j}{|\zeta|^3} d^3\zeta = \alpha_{jk} y_k +$$

$$+ \frac{3}{4\pi} \int_{(V_\infty)} \frac{\zeta_j}{|\zeta|^3} d^3\zeta = \frac{3}{4\pi} \int_{(V_\infty)} \frac{\alpha_{jk} y_k - \zeta_j}{|\alpha y - \zeta|^3} d^3\zeta, \tag{1.38}$$

by utilizing formulae (1.13) and (1.35). On the other hand, it is possible to prove the validity of the correlation

$$\frac{3}{4\pi} \int_{(V_\infty)} \frac{\alpha_{jk} y_k - \zeta_j}{|\alpha y - \zeta|^3} d^3\zeta \equiv \alpha_{jk} \frac{3}{4\pi} \int_{(V_\infty)} \frac{y_k - \zeta_k}{|y - \zeta|^3} d^3\zeta +$$

$$+ (\alpha_{jk} - \delta_{jk}) \int_{(V_\infty)} \frac{\zeta_k}{|\zeta|^3} d^3\zeta. \tag{1.39}$$

This proof is based on the definition that the integrals are identical if their difference is a convergent integral identical to zero. Taking into consideration formulae (1.37) and (1.38), we will rewrite formula (1.39) as follows:

$$x_j = \alpha_{jk} Y_k + (\alpha_{jk} - \delta_{jk}) \int_{(V_\infty)} \frac{\zeta_k}{|\zeta|^3} d^3\zeta. \tag{1.40}$$

The latter appears to be a formula for the reorientation of coordinates in generalized integrals, which corresponds to the rotation of the coordinate axis. The reorientation formulae, which are opposite to formulae (1.40), are the following:

$$Y_j = \beta_{jk} X_k + (\beta_{jk} - \delta_{jk}) \int_{(V_\infty)} \frac{\zeta_k}{|\zeta|^3} d^3\zeta. \tag{1.41}$$

The reorientation of the x_j coordinates, which correspond to the shift of their origin without rotating the coordinate axes, in accordance with formulae

$$x_j = y_j - y_j^{(0)}, \quad y_j = x_j + x_j^{(0)}, \tag{1.42}$$

correspond in the system of coordinates (expressed in generalized integrals) to the reorientations

$$X_j = Y_j - y_j^{(0)} = \frac{3}{4\pi} \int\limits_{(V_\infty)} \frac{y_j - \zeta_j}{|y - \zeta|^3} d^3\zeta - y_j^{(0)} =$$

$$= \frac{3}{4\pi} \int\limits_{(V_\infty)} \frac{y_j - y_j^{(0)} - \zeta_j}{|y - y^{(0)} - \zeta|^3} d^3\zeta, \tag{1.43}$$

$$Y_j = X_j + y_j^{(0)} = \frac{3}{4\pi} \int\limits_{(V_\infty)} \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta + y_j^{(0)} =$$

$$= \frac{3}{4\pi} \int\limits_{(V_\infty)} \frac{x_j + y_j^{(0)} - \zeta_j}{|x + y^{(0)} - \zeta|^3} d^3\zeta.$$

We shall not discuss in detail the problem of reorientation of the system of coordinates X_j which are expressed in generalized integrals. We will only mention that formulae (1.40) and (1.41) can be rather easily generalized in the case of a conversion of the right angles into oblique angles. It should also be mentioned that the mutually defined conformity (1.34) does not appear to be an isomorphism, since the linear combination of the integrals X_j (only the difference $X_j - X_j'$ corresponds to the difference $x_j - x_j'$) does not correspond to the linear combination of the numbers x_j. The absence of isomorphism is also

expressed by the fact that the reorientation of coordinates x_j expressed in common values in accordance with formula (1.35) does not correspond to the conversion

$$X_j = \alpha_{jk} Y_k,$$

/75

which should have been anticipated if isomorphism had been observed. An additional term is contained in (1.40)

$$(\alpha_{jk} - \delta_{jk}) \int\limits_{(V_\infty)} \frac{\zeta_j}{|\zeta|^3} d^3\zeta,$$

the presence of which is stipulated by the absence of the above mentioned isomorphism.

Let us assume that the assigned invariant function Φ of the coordinates x_j is

$$\Phi = \Phi\ (x). \qquad (1.44)$$

If we should substitute the coordinate x_j with its value:

$$x_j = \frac{3}{4\pi} \int\limits_{(V_\infty)} \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta + \frac{3}{4\pi} \int\limits_{(V_\infty)} \frac{\zeta_j}{|\zeta|^3} d^3\zeta = X_j - X_j^{(0)}, \qquad (1.45)$$

then we obtain

$$\Phi = \Phi\ (x) = \Phi\ (X - X^0). \qquad (.146)$$

Thus, we have defined the function Φ (X), the arguments of which are the coordinates X_j of the spatial points, expressed in generalized integrals. The expression for the distance of two points in Euclidean space can serve as an example:

$$r = |x - x'| = |X - X'|.$$

Another example is the density function of matter in universal space

$$\rho = \rho\ (x),$$

which we consider as an invariant function in relation to the reorientation of spatial coordinates, as well as an invariant in relation to the displacement of the origin of the coordinates. In accordance with formula (1.46) we obtain:

$$\rho\ (x) = \rho\ (X - X^0).$$

We will always consider the invariant function of the coordinates of spatial point, as being expressed in actual values, regardless of whether the coordinates are given in actual values or in generalized integrals.

§2. The Poisson Equation and the Equations of Motion of a Test Particle in the Field of Universal Gravitation /76

Newton's theory of gravity is based primarily on two groups of equations. The first group can be found from the distribution of gravitating masses of a field and is known as field equations. In a non-relativistic approximation, such an equation is the Poisson equation in which the field is defined by the potential Ω. The second group of equations determines the motion of the point masses in the given field and is known as the equations of motion.

In accordance with the above statement, the theory of universal gravitation, which takes into consideration the presence of masses in the entire, infinite universe, must also be broken down into divisions containing field equations and the equations of motion of a point mass in the field of universal gravitation.

The Poisson equation expresses the law of gravity in an infinitesimal volume of space and combines the derivatives of the potential Ω according to the coordinates of the spatial points with the density of matter within that circumference. The presence of foreign gravitating masses outside of the small space which is being examined does not enter into the Poisson equation. The gradient of the potential is a measurable magnitude (the force of gravity per unit of mass), and it is therefore possible to verify experimentally the law of gravity, or the Poisson equation. In a verification it suffices to have the data

on the matter distribution in a volume of space which is being investigated and it is not expedient to consider the presence of matter outside of this space. We consider that an experimental check confirms fully the accuracy of the Poisson equation, if only the discussion pertains to a non-relativistic approximation.

The difficulties related to the law of gravity occur during the integration of the Poisson equation for an infinite volume. Should we wish to obtain the potential Ω in the form of a definite function of coordinates, and represent Ω in actual numbers, we shall be confronted by a gravitational paradox.

In the preceding paragraph of this paper it was shown that through a generalization of the concept of the integral, it is possible to solve the Poisson equation without establishing the boundary conditions by means of the integral

$$\Omega = \varkappa \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{d^3\zeta}{|x - \zeta|}\ , \tag{2.1}$$

which does not converge but has a definite mathematical value. Of course, Ω according to formula (2.1) is not a common function in actual numbers. However, if we succeed in proving that it is possible by means of a generalized integral Ω, to solve the problem of the motion of a point mass in the field of universal gravitation and to calculate the particle's trajectory, we shall have the entire foundation for considering the function Ω which is determined by formula (2.1), which fully merits the title of potential. /77

The motion of a test particle, i.e., of an infinitesimal point mass in the field of gravity according to Newton's theory,is described by the equation

$$\frac{d^2x_j}{dt^2} = \frac{\partial}{\partial x_j}\left\{\varkappa \int\limits_{(V)} \rho\ (\zeta)\ \frac{d^3\zeta}{|x - \zeta|}\right\} = \varkappa \int\limits_{(V)} \rho\ (\zeta)\ \frac{\zeta_j - x_j}{|x - \zeta|^3}\ d^3\zeta, \tag{2.2}$$

with the masses creating the field and having a density $\rho\ (\zeta)$. These masses are found in a definite finite volume V of space. The presence of masses not included in volume V is not taken into consideration.

In the real conditions of equation (2.2) the exact motion of the particle is not described since the foreign masses are disregarded. It

is possible to dispose of the error which we obtain here by augmenting the volume of the integration V so that it will encompass the previously excluded masses. Finally, this leads to a limitless increase of the volume V. We obtain the equation

$$\frac{d^2x_j}{dt^2} = \lim_{V\to\infty}\left\{\varkappa \int_{(V)} \rho\ (\zeta)\ \frac{\zeta_j - x_j}{|\zeta - x|^3}\, d^3\zeta\right\} = \varkappa \int_{\infty} \rho\ (\zeta)\ \frac{\zeta_j - x_j}{|\zeta - x|^3}\, d^3\zeta. \qquad (2.3)$$

But the objective is still not accomplished since the right hand terms of equation (2.3) diverge. The entry (2.1) becomes a paradox since the divergent right term is united with the convergent left term by the equal sign.

Returning to an examination of the equations (2.2) we note that their right terms are gradients of the potential

$$\Omega = \varkappa \int_{(V)} \rho\ (\zeta)\ \frac{d^3\zeta}{|x - \zeta|},$$

where the integral is obtained in accordance with the finite volume V. The right terms of equation (2.3) as boundaries where V approaches infinity have no meaning since such a boundary does not exist. However, they have a mathematical meaning if the existent integrals are considered as integrals with an actually infinite volume of integration V_∞. The right terms of the equations

$$\varkappa \int_{(V_\infty)} \rho\ (\zeta)\ \frac{\zeta_j - x_j}{|\zeta - x|^3}\, d^3\zeta \qquad (2.4)$$

are, then, gradients of the potential

$$\Omega = \varkappa \int_{(V_\infty)} \rho\ (\zeta)\ \frac{d^3\zeta}{|\zeta - x|}, \qquad (2.5)$$

whereby the potential Ω satisfies the Poisson equation according to the results of the preceding paragraph.

The substitution of the right terms of equation (2.3) by integrals (2.5) also requires certain substitutions of their left terms. Indeed, integrals (2.5) are not expressed in numbers. Hence, the left terms of the equations should also be generalized integrals and not common functions

$$\frac{d^2 x_j}{dt^2}$$

of the variable t.

In the preceding paragraph, it was proved that the coordinates of the spatial points can be expressed in generalized integrals of an appropriate type. With this possibility, we assume that the coordinates of the test particle have the form

$$\xi_j = \frac{3}{4\pi} \int\limits_{(V_\infty)} a\ (t)\ \frac{x_j\ (t) - \zeta_j}{|x\ (t) - \zeta|^3}\, d^3\zeta\ , \tag{2.6}$$

where the density function a (t) is independent of the variables of integration ζ_j but is dependent on time t. The arguments of the integral x_j (t), which are the coordinates of the test particles in actual numbers, are also dependent on t so that in the general calculation

$$\xi_j = \xi_j\ (t).$$

The dependence of ξ_j on t expresses the motion of the particle. On the basis of the stated assumptions, the equation of motion of the test particle should be written as follows

$$\frac{d^2\xi_j}{dt^2} = \varkappa \int\limits_{(V_\infty)} \rho\ (\zeta)\ \frac{\zeta_j - x_j}{|\zeta - x|^3}\, d^3\zeta = \frac{\partial\, \Omega}{\partial x_j}\ , \tag{2.7}$$

$$\xi_j = \frac{3}{4\pi} \int\limits_{(V_\infty)} a\ (t)\ \frac{x_j - \zeta_j}{|\zeta - x|^3}\, d^3\zeta \tag{2.8}$$

and the equations are free from contradiction.

If equation (2.7) together with equation (2.8) are actually the equations of motion of the test particle, they must have solutions describing the actual motion of the particle. We will prove that this is so.

Utilizing formula (1.28) we combine equations (2.7) and (2.8) into one equation

$$a \frac{d^2x_j}{dt^2} + 2 \frac{da}{dt} \frac{dx_j}{dt} = \varkappa \int_{(V_\infty)} \left[\rho\,(\zeta) + \frac{3}{4\pi} \frac{1}{\varkappa} \frac{d^2a}{dt^2}\right] \frac{\zeta_j - x_j}{|\zeta - x|^3}\, d^3\zeta. \qquad (2.9)$$

We utilize thereby the rules of mathematical operations with generalized integrals. In these equations the left terms are common functions of time t and, consequently, the right terms should converge. We will assume that a mean density $\overline{\rho}$ of matter exists in infinite space in the sense of formula (1.17). Then the integral /79

$$\int_{(V_\infty)} (\rho - \overline{\rho}) \frac{\zeta_j - x_j}{|\zeta - x|^3}\, d^3\zeta = \int_{(\infty)} (\rho - \overline{\rho}) \frac{\zeta_j - x_j}{|\zeta - x|^3}\, d^3\zeta$$

converges and belongs to the convergent integrals subclass of the general class of integrals. By selecting in the subintegral expression of the right term

$$\frac{3}{4\pi} \frac{1}{\varkappa} \frac{d^2a}{dt^2} = -\overline{\rho}\,, \qquad (2.10)$$

we obtain

$$a \frac{d^2x_j}{dt^2} + 2 \frac{da}{dt} \frac{dx_j}{dt} = \varkappa \int_{(\infty)} (\rho - \overline{\rho}) \frac{\zeta_j - x_j}{|\zeta - x|^3}\, d^3\zeta, \qquad (2.11)$$

the right and left terms of which are functions in actual numbers.

A general solution to equation (2.10) is

$$a\,(t) = A + Bt - \frac{1}{2} \Lambda^2 t^2, \qquad (2.12)$$

where A and B are constants of integration and

$$\Lambda^2 = \frac{4\pi}{3} \varkappa \bar{\rho} \ . \tag{2.13}$$

The coordinates ξ_j of the test particle in motion, are obtained, according to formula (2.6), in the form of generalized integrals

$$\xi_j = \frac{3}{4\pi} \int\limits_{(V_\infty)} (A + Bt - \frac{1}{2} \Lambda^2 t^2) \frac{\dot{x}_j (t) - \zeta_j}{|x - \zeta|^3} d^3\zeta, \tag{2.14}$$

whereby x_j (t) are the solutions to equations (2.11) which, after substitution in place of a (t), of its value according to formula (2.12), adopt the following form

$$(A + Bt - \frac{1}{2} \Lambda^2 t^2) \frac{d^2 x_j}{dt^2} + 2 (B - \Lambda^2 t) \frac{dx_j}{dt} =$$

$$= \varkappa \int\limits_{(V_\infty)} (\rho - \bar{\rho}) \frac{\zeta_j - x_j}{|x - \zeta|^3} d^3\zeta. \tag{2.15}$$

We will mention that x_j (t) are also coordinates of the /80
particle in motion; however, contrary to the coordinates ξ_j (t) these are expressed in actual numbers.

And so we have proved that equation (2.7) together with equation (2.8) have solutions, whereby the problem of their solution is reduced to the solving of the common differential equation (2.15). The generalized integrals of the form in (2.14) are interpreted as coordinates of the test particle in motion. The function of time

$$x_j (t)$$

is contained in the subintegral expression of the integral (2.14), which is the solution to equation (2.15), which are also the coordinates of

the particles in motion, but expressed here by common numbers. Finally, we have two systems of coordinates, the sustems ξ_j and x_j, in which motion is considered. The relation between them is expressed by formula (2.14).

Let us assume that $x_j'(t)$ and $x_j''(t)$ are two particular solutions to the equation (2.15) corresponding to the two concrete test particles. According to formula (1.33), the distance R between these particles is expressed in the following way:

$$R = |\xi' - \xi''| = (A + Bt - \frac{1}{2}\Lambda^2 t^2)\ |x' - x''| =$$

$$= (A + Bt - \frac{1}{2}\Lambda^2 t^2)\ r, \tag{2.16}$$

where

$$r = |x' - x''|\ .$$

From formula (2.16) it follows that the systems of coordinates ξ_j and x_j correspond to the various scales of the distance measurements, whereby the correlation of the scales changes with time t. We can standardize the solutions to the equation (2.10) so that the distance scales coincide when the time $t = 0$.

In order to do this, it is necessary to consider that

$$A = 1,$$

then the standardized solution to equation (2.10) will be

$$a(t) = 1 + Bt - \frac{1}{2}\Lambda^2 t^2, \tag{2.17}$$

in which the constant B remains undetermined. As far as the constant A^2 is concerned it is combined, according to formula (2.13), with the mean density of matter in the infinite universe and can be considered as a certain universal constant. We shall assume that B in formula (2.17) is a universal constant whose magnitude we can obtain by appropriate /81 measurements.

Equation (2.7) or their equivalent equations of motion (2.15) have the accuracy of a non-relativistic approximation. The application of non-relativistic correlations in cosmology is justified if the universe is in a quasi-stationary state or if the matter density ρ changes with time so slightly that the dependence of the integral

$$\int_{(\infty)} (\rho - \bar{\rho}) \frac{\zeta_j - x_j}{|x - \zeta|^3} d^3\zeta \tag{2.18}$$

on time t is not affected by the change of density ρ in the subintegral expression. If this is so, then for the computation of the integral (2.18) it suffices to have the data concerning the matter distribution in a certain large volume V, since from the convergence of the integral (2.18) it follows that

$$\int_{(\infty)} (\rho - \bar{\rho}) \frac{\zeta_j - x_j}{|\zeta - x|^3} d^3\zeta \approx \int_{(V)} (\rho - \bar{\rho}) \frac{\zeta_j - x_j^{\,3}}{|\zeta - x|} d^3\zeta. \tag{2.19}$$

Hereby, it is assumed that the investigated test particle with the coordinates x_j (t) is located within volume V' situated in the center of volume V and that

$$V' \ll V. \tag{2.20}$$

Thus, in order to compute the motion of the test particle, the data concerning the matter distribution in the finite universe as a whole are not essential. It suffices to have concrete values of the density function $\rho(\zeta)$ in a large but finite volume V. However, the motion of the particles can be calculated if they are located within volume V', which in turn is located in the central part of volume V, whereby V' should be much less than V. The applicability of the equations of motion (2.15) in time is also limited in the last case. The solutions to equation (2.15) are of an objective value only for the time interval in which the test particle leaves the volume V'.

The fact that the equations of motion of the test particle described with sufficient accuracy the motions within the time interval $-t_0$ and $+t_0$ and within volume V' only, does not constitute a deficiency of the theory. In particular it is not solely characteristic for a non-relativistic approximation. Actually, if a theory that described

the motion and the conditions of matter in the time interval from $-\infty$ to $+\infty$ and throughout the universe, would have been available, such a theory would have solved problems which cannot be solved by experiments or observations. Such a theory cannot be an acknowledged physical theory. The inaccuracy of the equations of motion (2.15) of a test particle /82 in the described sense, constitutes the principal inaccuracy, which is inherent to reasonable cosmology that acknowledges the infinity of the universe.

§3. Another Presentation of the Equations of Motion of a Test Particle in a Field of Universal Gravitation

It was proved in the preceding section that the equations of motion of a test particle consist of the differential equation (2.7) and the equation (2.9), which combines the coordinates x_j of the test particle in actual numbers, with the coordinates ξ_j in generalized integrals. The exclusion of ξ_j from equations (2.8) leads to equation (2.15) in which all the coordinates are represented in actual numbers. However, it is of interest to find equations of motion in which the coordinates are represented in generalized integrals only. Equations of such a type will be more symmetrical than equation (2.15) and, obviously, also more consistent from the viewpoint of the concepts which were developed in this paper.

The formulae of the reorientation which combines the system of coordinates ξ_j in generalized integrals, with the system of coordinates x_j in common numbers, are defined in accordance with formula (2.8) as follows:

$$\xi_j = \frac{3}{4\pi} \int\limits_{(V_\infty)} a\ (t)\ \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta \tag{3.1}$$

and in accordance with (2.7) the equations of motion are as follows

$$\frac{d^2\xi_j}{dt^2} = \varkappa \int\limits_{(V_\infty)} \rho\ (x')\ \frac{x'_j - x_j}{|x' - x|^3}\, d^3x'. \tag{3.2}$$

The coordinate of the spatial point of integration is the variable of the integration x'_j of the right term of equation (3.2). We shall determine:

$$X_j = \frac{3}{4\pi} \int_{(V_\infty)} a\ (t)\ \frac{x'_j - \zeta_j}{|x' - \zeta|^3}\, d^3\zeta. \tag{3.3}$$

Evidently, X_j is the coordinate of the spatial point of integration of the integral (3.2) in generalized integrals, and equation (3.3) combines with x'_j with X_j. Since /83

$$X_j - \xi_j = \frac{3}{4\pi} \int_{(V_\infty)} a\ (t)\ \frac{x'_j - \zeta_j}{|x' - \zeta|^3}\, d^3\zeta -$$

$$- \frac{3}{4\pi} \int_{(V_\infty)} a\ (t)\ \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta = a\ (x'_j - x_j), \tag{3.4}$$

then the expression

$$\frac{X_j - \xi_j}{|X - \xi|^3}$$

has the value of an ordinary number:

$$\frac{X_j - \xi_j}{|X - \xi|^3} = \frac{1}{a^2}\, \frac{x'_j - x_j}{|x' - x|^3}\ .$$

Furthermore, by determining

$$dX_j = \frac{3}{4\pi} \int_{(V_\infty)} \frac{x'_j + dx'_j - \zeta_j}{|x' + dx' - \zeta|^3}\, d^3\zeta - \frac{3}{4\pi} \int_{(V_\infty)} \frac{x'_j - \zeta_j}{|x' - \zeta|^3}\, d^3\zeta = a \cdot dx'_j\ ,$$

it is possible to write

$$d^3x' = \frac{1}{a^3} d^3X,$$

then d^3X is also expressed by an actual number. The density function $\rho(x')$ in the right part of equation (3.2) is the invariant (in a non-relativistic approximation). According to the formula

$$\rho\ (x') = \rho\ (X - X^0),$$

we consider the matter density as a function of coordinates X_j of the integration space, whereby the coordinates X_j are determined by formula (3.3). Finally, we find that the symbol

$$\int\limits_{(V_\infty)} \rho\ (X) \frac{X_j - \xi_j}{|X - \xi|^3} d^3X$$

is in essence a generalized integral, since the subintegral expression is represented in actual numbers. Utilizing this circumstance, we write the equation of motion of the test particle in a field of universal /84 gravitation in the following way:

$$\frac{d^2\xi_j}{d\tau^2} = \varkappa \int\limits_{(V_\infty)} \rho\ (X) \frac{X_j - \xi_j}{|X - \xi|^3} d^3X, \tag{3.5}$$

where all the coordinates are presented by the generalized integrals X_j and ξ_j. We will consider equation (3.5) as a generalization of the equation

$$\frac{d^2x_j}{dt^2} = \varkappa \int\limits_{(V)} \rho\ (x') \frac{x_j' - x_j}{|x' - x|^3} d^3x',$$

which describe the motion of a test particle in the field of a finite system of gravitating masses, which are included in the finite volume V

of space. In equation (3.5) the time variable is designated by the letter τ. The substitution of t by τ meanwhile is only a formal change of designation. However, as we will see later on, t and τ have a distinct physical meaning.

We shall prove that the equation (3.5) have the following solutions

$$\xi_j = \frac{3}{4\pi} \int\limits_{(V_\infty)} a\,(t)\, \frac{x_j\,(t) - \zeta_j}{|x - \zeta|^3}\, d^3\zeta\ , \tag{3.6}$$

where x_j (t) as a common function of the variable t satisfies the equation

$$\frac{d^2 x_j}{dt^2} = G \int\limits_{(\infty)} (\rho - \bar{\rho})\, \frac{\zeta_j - x_j}{|\zeta - x|^3}\, d^3\zeta . \tag{3.7}$$

The time variant τ is a certain function of variable t:

$$\tau = \tau\,(t) \tag{3.8}$$

and the density function a(t) in (3.6) satisfies the equation

$$\frac{d^2 a}{dt^2} - \frac{2}{a}\left(\frac{da}{dt}\right)^2 + a\,\frac{4\pi G}{3}\,\bar{\rho} = 0 . \tag{3.9}$$

It is also assumed that the integrals of the right terms of equation (3.7) converge or that the average density of matter $\bar{\rho}$ exists in the infinite universe according to the definition in (1.17).

Proceeding with the proof of this assertion, we differentiate the correlations (3.6) twice by the variable t:

$$\frac{d\xi_j}{dt} = \frac{3}{4\pi} \int\limits_{(V_\infty)} \frac{da}{dt}\, \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta + a\,\frac{dx_j}{dt}\ ,$$

$$\frac{d^2\xi_j}{dt^2} = \frac{3}{4\pi} \int\limits_{(V_\infty)} \frac{d^2 a}{dt^2}\, \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta + 2\,\frac{da}{dt}\,\frac{dx_j}{dt} + a\,\frac{d^2 x_j}{dt^2} .$$ /85

Herefrom, after simple conversions we obtain

$$\frac{d^2\xi_j}{dt^2} - \frac{2}{a}\frac{da}{dt} \cdot \frac{d\xi_j}{dt} = \frac{3}{4\pi} \int\limits_{(V_\infty)} \left(\frac{d^2a}{dt^2} - \frac{2}{a}\left(\frac{da}{dt}\right)^2 \right) \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta + a \frac{d^2x_j}{dt^2} . \tag{3.10}$$

Inasmuch as x_j (t) satisfies the equation (3.7), then by substitution

$$\frac{d^2x_j}{dt^2}$$

in the equation (3.10) with their values from the equation (3.7) we find, after simple calculations

$$\frac{d^2\xi_j}{dt^2} - \frac{2}{a}\frac{da}{dt}\frac{d\xi_j}{dt} = G \int\limits_{(V_\infty)} a\rho\ (\zeta) \frac{\zeta_j - x_j}{|\zeta - x|^3} d^3\zeta +$$

$$+ \frac{3}{4\pi} \int\limits_{(V_\infty)} \left(\frac{d^2a}{dt^2} - \frac{2}{a}\left(\frac{da}{dt}\right)^2 + a\frac{4\pi G}{3}\bar{\rho} \right) \frac{x_j - \zeta_j}{|x - \zeta|^3} d^3\zeta. \tag{3.11}$$

We shall choose the function a (t) to satisfy the equations

$$\frac{d^2a}{dt^2} - \frac{2}{a}\left(\frac{da}{dt}\right)^2 + a\Lambda^2 = 0, \tag{3.12}$$

$$\Lambda^2 = \frac{4\pi G}{3}\bar{\rho} . \tag{3.13}$$

Then in place of (3.11) we obtain

$$\frac{d^2\xi_j}{dt^2} - \frac{2}{a}\frac{da}{dt}\frac{d\xi_j}{dt} = G \int a\rho\ (x') \frac{x'_j - x_j}{|x' - x|^3} d^3x', \tag{3.14}$$

where the variable of the integration ζ_j in the right parts of equation (3.14) is substituted by the variable x'_j (this is only a formal substitution). In the subintegral expressions of the right terms of equation (3.14), the variables x_j and x'_j are the coordinates of the spatial points in actual numbers. Applying the conversions (3.1) and (3.3) we can write

$$a\rho\ (x')\ \frac{x'_j - x_j}{|x' - x|^3}\ d^3x' = \rho\ (X)\ \frac{X_j - \xi_j}{|X - \xi|^3}\ d^3X\ ,$$

so that equation (3.14) adapts the form /86

$$\frac{d^2\xi_j}{dt^2} - \frac{2}{a}\frac{da}{dt}\frac{d\xi_j}{dt} = G \int\limits_{(V_\infty)} \rho\ (X)\ \frac{X_j - \xi_j}{|X - \xi|^3}\ d^3X. \tag{3.15}$$

In order to obtain the equation (3.15) from the latter, we introduce in place of the time variable t, a new variable

$$\tau = \tau\ (t), \tag{3.16}$$

which satisfies the equation

$$\frac{d^2\tau}{dt^2} - \frac{2}{a}\frac{da}{dt}\frac{d\tau}{dt} = 0, \tag{3.17}$$

where a (t) is the solution to equation (3.12). Then from equation (3.15) we obtain

$$\frac{d^2\xi_j}{d\tau^2} = \varkappa \int\limits_{(V_\infty)} \rho\ (X)\ \frac{X_j - \xi_j}{|X - \xi|^3}\ d^3X, \tag{3.18}$$

where

$$\varkappa = \left(\frac{dt}{d\tau}\right)^2 G. \tag{3.19}$$

Finally, we arrive at the fact that, for the motion of the test particle, there are two types of equations - equations (3.18) and (3.17). In Equation (3.18) the coordinates of the test particle and the space coordinates of the integration are expressed by generalized integrals. In equation (3.7) the very same coordinates are represented by actual numbers. In addition, there are two distinct variables of time, t and τ, whereby t is included in equation (3.7) and τ in equation (3.18). The space-time continuum, on the basis of which the motion of the test particles was studied, is represented in the first case by a system of coordinates

$$(x,\ t), \tag{3.20}$$

and in the second case, by the following system of coordinates

$$(\xi,\ \tau), \tag{3.21}$$

whereby the space coordinates ξ_j and x_j are combined by the correlation (3.6) and the time variables τ and t are combined by the correlation (3.16).

In the preceding section of this paper, the equations of motion of the test particle were written

$$\frac{d^2\xi_j}{d\tau^2} = \varkappa \int_{(V_\infty)} \rho\ (x') \frac{x'_j - x_j}{|x' - x|^3}\, d^3x', \tag{3.22}$$

whereby /87

$$\xi_j = \frac{3}{4\pi} \int_{(V_\infty)} a\ (\tau) \frac{x_j - \zeta_j}{|x - \zeta|^3}\, d^3\zeta \tag{3.23}$$

and the function a (τ) satisfies the equations*

$$\frac{d^2a}{d\tau^2} = -\Lambda^2, \quad \Lambda^2 = \frac{4\pi\varkappa}{3}\rho. \tag{3.24}$$

---

*See equations (2.7), (2.8), and (2.10). We substitute for the variable t in these equations the variable τ, inasmuch as the variable t in Section 2 coincides with the variable τ in Section 3.

On the other hand, from this section we see that the function a (τ) must satisfy equation (3.12).

After the substitution of the variable t by the variable τ, equation (3.12) acquires the form

$$\frac{d^2a}{d\tau^2} = - \Lambda^2 \frac{1}{a^3} . \tag{3.25}$$

Actually, the first integral of the differential equation (3.17) is

$$\frac{d\tau}{dt} = C\ a^2, \tag{3.26}$$

where C is the constant of integration, the concrete value of which is determined by the requirement

$$\left(\frac{d\tau}{dt} \right)_{t=\tau=0} = 1, \tag{3.27}$$

so that

$$C = \left(\frac{1}{a^2} \right)_{t=\tau=0} . \tag{3.28}$$

For an explanation of the physical meaning of equation (3.27) we assume the existence of two types of clocks which measure time t and time τ. Inasmuch as $d\tau/dt$ is not a constant, τ- and t-clocks cannot be synchronized for a finite interval of time. But they can be synchronized for an infinitely short interval of time which contains the moment t = 0. Considering the moment t = 0 as a moment of present time, we regulate the t- and τ-clocks so that their units (seconds) coincide. Then equation (3.27) will be satisfied. We shall call this a condition of the relative calibration of t- and τ-clocks.

Let us assume that there are spatial points A and A', the coordinates of which are ξ_j and ξ'_j in the system (ξ) and, respectively, x_j and x'_j in the system (x). We have

$$|\xi - \xi| = a\ (\tau)\ |x - x'|. \tag{3.29}$$

We can, although until now only formally, consider the magnitude /88
$|x - x'|$ as the distance between points in the system (x) and $|\xi - \xi'|$ as the distance in the system (ξ). In addition, it is possible to require that in the moment of time t = 0 both distances coincide, or according to (3.29)

$$(a\ (\tau))_{t=\tau=0} = 1. \tag{3.30}$$

We shall call this the condition of relative calibration in the scale of the distance measurement in the (ξ)- and (x)-systems. From equation (3.28) it follows that C = 1 and on the basis of equation (3.26), we have

$$\frac{d\tau}{dt} = a^2. \tag{3.31}$$

Utilizing the last expression, it is possible to obtain equation (3.25) by means of the conversion of equation (3.12).

Equations (3.24) and (3.25) do not accurately coincide, and therefore it is possible to prove that we have two different theories: the theories of the preceding and present sections. However, this difference is only formal. In reality the solution to the equation in the form of a power series, and considering the calculation of the quadratic terms, is

$$a\ (\tau) = 1 + B\tau - \frac{1}{2}\Lambda^2\tau^2 + \cdots, \tag{3.32}$$

where B is the constant of integration, and a (τ) satisfies the original conditions (3.30). But the power series (3.32) with its terms written out also satisfies the equation (3.24) so that the difference between the solutions of equations (3.24) and (3.25) begins with the calculations of the numbers of the third and higher orders of series (3.32). In the interval

$$-\tau_0 \ldots + \tau_0, \tag{3.33}$$

if τ_0 is not too great, the differences between the solutions of the cited equations practically do not exist, and they are both equally suitable for a description of the test particle in the field of universal gravitation. In connection with the latter, it is necessary to mention

that on the basis of the material described in Section 2 of this paper, from equations of motion, it is generally impossible to require that they describe the motion of the test particle over any given interval of time. We consider that the interval of time in (3.33) is the exact time necessary for meaningful solutions to equation (3.24) or (3.25).

On the basis of the stated equation [(3.5) or (3.22)] the equations for the determination of motion of the test particle can be considered to be accurate, and the question concerning which is to be given preference is a matter of taste. However, in view of the compactness and /89 complete analogy in the equations of motion of the test particle under the effect of a finite system of gravitating masses, the author of this paper prefers equation (3.5) to equation (3.22).

In equation (3.5) the coordinates ξ_j of the test particle are represented in generalized integrals, but the time variable τ is expressed by an ordinary number. A definite asymmetry in the represented coordinates of the space-time continuum evidently is effected by a non-relativistic solution of the problem.

Although the exact solution of equation (3.5) for a suitably long interval of time does not have a satisfactory physical value, as an example of an accurate solution, which is given by the following expressions for a (t) and τ (t), it can introduce definite interests:

$$a\,(t) = e^{\pm\Lambda t}, \tag{3.34}$$

$$\tau = \pm\,\frac{1}{2\Lambda}\left(e^{\pm 2\Lambda t} - 1\right). \tag{3.35}$$

In his time, Milne proposed a hypothesis according to which two systems of time t and τ exist. In them, τ is cosmological time and t is conventional time. The relation between τ and t according to Milne is conveyed by the formula (3.35), and therefore, we shall call expressions (3.34) and (3.35) Milne's solutions.

§4. Concerning the Existence of Two Systems of Measurement of Space and Time

Equations (3.5) are equations of motion of a test particle in (ξ, τ)-system of space and time coordinates. With a conversion by means of formula (3.6) from equation (3.5), we obtain equation (3.7), which also describe the motion of a test particle; however, not in the (ξ, τ)-

system but in the (x, t)-system of coordinates. Finally, there are two types of equations of motion of a particle, equations (3.5) and (3.7); and there are, respectively, two systems of coordinates of space and time, (ξ, τ)- and (x, t)-systems.

If the integrals ξ_j, according to formula (3.6), are the coordinates of the spatial points, then the difference between the coordinates is an actual number. By defining

$$r_{AA'} = |\xi - \xi'| = a\ (t)\ |x - x'| \tag{4.1}$$

as the distance between the points with the coordinates ξ_j and ξ'_j, the distance as a measurable magnitude will be expressed in actual numbers. The distance between the same points is calculated by means of their coordinates x_j and x'_j.

$$R_{AA'} = |x - x'|. \tag{4.2}$$

Therefore, formulae (4.1) and (4.2) yield /90

$$r_{AA'} = a\ (t)\ R_{AA'}\ , \tag{4.3}$$

whereby the function a (t) is the solution to equation (3.9).

In (ξ, τ)-system the time lapse is expressed by the variability of the values of τ and, respectively, in the (x, t)-system by the variability of the value t, whereby τ and t are combined by the functional correlation

$$\tau = \tau\ (t), \tag{4.4}$$

and τ (t) satisfies the differential equation (3.31).

The utilization of the two qualitatively distinct systems of coordinates rendered it possible to remove from the theory of universal gravitation the contradiction in the form of a gravitational paradox. However, the solution to the problem has so far a formal character. For the acceptance of the physical theory, it is primarily necessary to indicate what formulae (4.3) and (4.4) designate and to disclose the physical essence of these formulae.

The distance between two points in space is a magnitude which can be obtained through measurement; for example, measurement by means of a rigid scale.

On the other hand, this distance can be obtained through a metric formula by calculating the coordinates of the points. We have not yet established which one of formulae (4.1) or (4.2) must be considered as a metric formula. An analogous problem also arises during an interpretation of formula (4.4). Time at a given point in space is measured in hours. The question consists of whether the hours show the variability of the values t or the values* of τ.

In order to obtain an answer to the posed questions we turn to the Weyl principle, which is the result of analyzing the basis for the general theory of relativity.

In 1921 Weyl indicated that, for the construction of a general theory of relativity, there is no need to assume the initial existence of instruments for the measurements of spatial distances and hours for the measurement of time. The instruments for the measurement of space-time distances are developed as a result of the theory and can be constructed on the basis of the theory which is already available. The general theory of relativity is primarily a theory of gravitation. According to the Weyl principle, the motion of matter (by the law of gravitation) determines the system of measurement of space-time distances. We shall call space and time measured by means of the gravitational /91 motion of matter, gravitational space and time.

For an explanation of the essence of the Weyl principle, it is interesting to examine the application of the verification of the passage of technical hours; that is, the instruments of time measurement which are based on a definite periodic motion of matter (for example, a pendulum clock) but as in any instrument, they do not measure a time lapse with complete accuracy. The passage of technical hours is checked by astronomical observations whereby the most accurate corrections are obtained from observations of the motion of the planets, the Moon and the Earth (the Sun). It is considered that the motion of members of the solar system complies with the differential equations of celestial mechanics which are obtained from the theory of gravitation. The equations

---

*In the present research, the problem of gravitation and also the problem of metrics of space and time, are examined in the non-relativistic approximation. Therefore, it is considered that space and time are divided: the first is measured by a rigid scale, the second by hours. However, we shall return to the problem of a rigid scale later.

and their solutions are considered to be elementary, and the passage of time is determined so that the calculated and observed positions of the planets coincide. It is evident that in the described method of measuring the passage of time there is the matter of a practical application of Weyl's principle and that the solar system represents a realization of the gravitational clocks.

In order to understand more fully the essence of the Weyl principle, we will note, that with the technological development of the application of time measurement in the last decade, atomic clocks, working on the basis of atomic processes, have been applied with great success. From among the atomic clocks, we shall examine those which work by means of a radioactive decomposition of atomic nuclei of certain chemical elements. Presently, the number of actually executed decompositions is distinguished from the number of nuclei potentially capable of such a process by means of counters. The rise in the number of impulses, which are registered by the counter, parallels the passage of time. However, the quantitative data concerning the passage of time is obtained from the law of radioactivity according to which the probability of disintegration per unit of time is independent of time. The increase in the number of decompositions is calculated and, according to an appropriate formula, the passage of time is determined on the basis of the counter readings.

It is evident that the law of radioactivity came first and that the determination of the passage of time in hours was developed as a result of this law. The circumstance, according to which the law of radioactivity is of a statistical character does not alter the nature of the method. In the described example of a clock, the Weyl principle moves into a region unforeseen by Weyl himself. Indeed, in order to formulate the law of radioactive decomposition, it is not expedient to assume the initial existence of a clock. The latter is developed on the basis of a law which has already been formulated, as was indicated above.

Having the gravitational and atomic clocks, it is possible to compare their motion. During a comparison, a preliminary relative calibration of these clocks is made. During a rather short time, the clocks are set up so that, on the average, their readings coincide with their motion. The shorter the duration of the relative calibration, the /92
more accurate is the result of this procedure. If the clocks measure time identically, the relative calibration is maintained forever. But inasmuch as gravitational clocks work on the basis of gravitational motion of matter, and atomic clocks work on the basis of atomic processes and inasmuch as the laws of macro and micro physics are qualitatively distinct, good reasons exist to assert the opposite: that atomic and gravitational clocks do not measure time in an identical way. It is also clear that both kinds of clocks are valid. The two systems of measuring

time are well founded, and both systems from a physical viewpoint are equally valid.

The general theory of relativity possesses completeness and independence. The completeness of the theory is expressed by the fact that the general theory of relativity exhaustively yields laws of gravitation and the motion of matter which is related to this theory. Its self-containment is reflected by the fact that the general theory of relativity does not require, for its structure or comprehension, any measuring instruments, whose physical essence is not encompassed by the theory. As far as the measurement of time is concerned, we have already tried to explain in the example of gravitational and atomic clocks, what the self-containment of the relativity theory consists of in relation to the measurement of time. The general theory of relativity enjoys an independence of exactly the same type as that of an instrument for the measurement of spatial distances. Returning to the example of the solar system, it is possible to establish the fact that this system not only realizes gravitational hours, but also determines with its own motion the scales for measurements of spatial distances. Not being able in the present article to discuss in detail the various constructions which realize the instruments for measurement of space and time in the general theory of relativity, we will mention that the construction of such a type has already been given in scientific literature (Pauli).

The completeness and self-containment of the general theory of relativity limits the area of its application. This theory clearly explains a defined range of physical phenomena but it cannot give an answer to the questions concerning, for example, the atomic phenomena. In connection with this, the discussion concerning the existence of the so-called rigid body has a certain significance. It has been asserted that the existence of a rigid body is not compatible with the principle of relativity. It seems to us that the last assertion can cause a misunderstanding. The rigid body exists, but the physical essence of the rigid body is not explained by the general theory of relativity.

Whether the rigid body exists or not does not concern the general theory of relativity; by the same token the theory is not concerned with the existence of radioactive elements. The rigid body is the material /93 for the creation of a solid core, which the instrument for measuring spatial distances represents. But, since the general theory of relativity does not require an axiom of the original existence of instruments for measuring space and time, and the solid core seems to be such an initially existing instrument in relation to the general theory of relativity, we again arrive at the conclusion that the general theory of relativity is completely indifferent to the problem of existence of a rigid body. However, if we attempt to identify the solid core with the instrument for distance measurement, obtained as a result of the general theory of

relativity, there arises a contradiction. In this sense, it is necessary to understand the assertion, according to which the rigid body is incompatible with the principle of relativity. Hence, if the rigid body exists, then we have in a physical sense, two distinct possibilities for measuring spatial distances. One of them applies the rigid core, the other applies to an appropriate instrument, constructed on the basis of gravitational motions of matter and defined by us as the gravitational scale of length.

If the general theory of relativity is complete and self-contained in the sense described above, then there should exist also other regions of physics with such a characteristic. We consider quantum mechanics or, in general, the physics of micro universe as another fully self-contained theory. Because of self-containment, the physics of the micro universe determines its own system of space and time measurement, which is known as the atomic system. Since the rigid body, as an idealization of a hard body, is the product of atomic forces and motions, the rigid core is obviously the instrument for distance measurement in an atomic system. Concurrent with the atomic clock, the rigid core pertains to a complex of instruments for space and time measurement within the atomic measurement system.

The hypothesis of the existence of two or more systems of space and time measurement constitutes, in essence, an expansion of the Weyl principle into various areas of physics, which encompass qualitatively distinct laws of motion of matter (macro and micro physics). Weyl himself attempted to utilize this principle for the creation of a single field theory, which would combine the entire physics in universal equations. However, none of the attempts in this direction have yielded satisfactory results. If we consider that the impossibility to create a single theory in the sense of Einstein and Weyl is a principal impossibility, then there is no single Weyl principle for physics as such. Physics is divided into individual, complete and self-contained branches, each of which has its own system of space and time measurement. The latter hypothesis does not contradict our concept concerning the unity of /94 space and time as a natural background for understanding of the essence of matter and its motion. Space and time are a uniform continuum of their points and in this continuity is contained the most general description that can be given on space and time. However, if we proceed with the measurement of the space-time continuum, then the results will be dependent on the physical essence of the measurement instruments.

A mathematical expression for the continuity of the space-time continuum permits the application of concepts of coordinates and coordinate systems. In principle, the coordinate system represents a number of abstract mathematical symbols with the power of a continuum

(actual numbers, generalized integrals of a special type, etc.), in which the continuum of the points of space-time are reflected. From these mathematical symbols the fulfillment of properties characteristic of the mathematical symbols are required. First of all, a potentiality of the determined mathematical operations is essential. But moreover in a majority of symbols, the meaning of the infinitesimal difference between two symbols must be determined. The mentioned quantity will be a system of coordinates if the corresponding symbols of two infinitely close points differ slightly.*

The potentiality for application of a large number of generalized integrals ξ_j of a coordinate system of three dimensional Euclidean space was utilized in the present work for the creation of the theory of universal gravitation in a non-relativistic approximation. This theory is free from a gravitational paradox. The analysis of the equations of motion of the test particle and the attempt to solve them indicates also the necessity to apply the systems of coordinates x_j in actual numbers. Finally, we have the systems of coordinates

$$(\xi, \tau) \text{ and } (x, t)$$

of space-time. These coordinates can be considered as belonging to distinct types. We consider that the distinction of the types reflects the distinctions of corresponding systems of space and time measurement, whereby the hypothesis of the existence of diverse systems of measurement should be considered a result of the general Weyl principle. Since, in the system (ξ, τ) the potential Ω and the equations of the test particle's motion are direct generalizations of the corresponding correlations of the classical theory of gravitation. This system of coordinates corresponds to the gravitational system of measurement. As far as the (x, t)-system is concerned there are no direct indications that it belongs to space and time measured by the rigid core and the atomic clocks. However, for lack of a better definition we preliminarily consider the (x, t)-system to be an atomic system of measurement. /95

The relation between gravitational time τ and atomic time t is expressed in formulae (3.31) and (3.32). The constant B in formula (3.32) expresses the acceleration of one time relative to the other:

$$\frac{d^2\tau}{dt^2} = B.$$

*The concept of infinitely close points comes from the idea of a continuum and does not propose a definite metric for space.

The universal constant B is, of course, a very small magnitude. Its experimental determination requires the comparison of the motion of atomic and gravitational clocks in the course of a prolonged interval of time. One can determine the constant B indirectly in the following way.

The evolution of the stellar system is regulated by the motion of matter under the influence of gravitational forces and takes place in gravitational time. The evolution of individual stars in the system is determined by the measurement of its inner structure primarily by the burning out of hydrogen, and therefore it takes place in atomic time. A comparison of the age of the stellar system calculated from dynamic factors and the age obtained from the change of the internal structure of a star makes it possible to determine the universal constant B. Another analogous opportunity for the indirect determination of the numerical value of the constant B consists of the comparison of the age of the solar system, which is calculated from the dynamic considerations, and the age obtained from the decomposition of uranium on the Earth.

The relation between the units of distance in (ξ)- and (x)-systems is given by formulae (3.29) and (3.32) from which it follows that one space expands relative to the other according to the law

$$\frac{d^2}{dt^2}\left(\frac{\Delta R}{\Delta r}\right) = 2\left(B^2 - \frac{1}{2}\Lambda^2\right), \tag{4.5}$$

if the system of time t is used, and according to the law

$$\frac{d^2}{d\tau^2}\left(\frac{\Delta R}{\Delta r}\right) = -\Lambda^2, \tag{4.6}$$

if the system of time τ is used. The well-known effect of the red shift of the spectral lines of distant spiral nebulae also indicates a certain expansion of the Metagalaxies. However, it is necessary to further develop the theory of red shift since it is still not clear whether equations (4.5) and (4.6) express this effect. [Equations (4.5) and (4.6) give a relative change of the scale of length; the red shift is a displacement of spectral lines.]

In conclusion it is still necessary to note that in the (x, t)-system, the vacuum has a gravitational characteristic. In fact, in the formula of the potential /96

$$\Omega = G \int_{(\infty)} (\rho - \tilde{\rho}) \frac{d^3\zeta}{|x - \zeta|} \tag{4.7}$$

the effective density is

$$\rho_{eff} = \rho - \tilde{\rho} .$$

The latter can be negative, and this takes place where the density of matter in the usual sense is equal to zero. From this follows the gravitational activity of a vacuum. In the (ξ, τ)-system a similar effect was not observed.

CONCERNING THE SOURCES OF THE INNER HEAT OF THE EARTH

By

E. A. Lubimova

The data have been examined concerning the role of short-lived radioactive isotopes in the thermal balance of the Earth, the release of gravitational and elastic energy, the generation of energy due to the changeover of the Earth, the generation of heat by tidal friction, the hypothesis on the role of the neutrino, and the new data on the natural radioactivity of long-lived isotopes. It has been indicated that a significant part of the gravitational energy released in the process of the condensation of the Earth can be transformed into the energy of elastic deformation. The activity of the short-lived isotopes can be significant only for the central part of the inner core of the Earth.

Recently, there have been assumptions made concerning the necessity to calculate in the thermal balance of the Earth a number of factors which we did not attempt to do earlier. Let us now examine the possible role of such factors as the energetic effect of short-lived isotopes, the emission of heat in the process of the gravitational changeover of the Earth, the hypothesis concerning the role of the neutrino, the emission of heat in the process of tidal friction, and also the new data concerning the mean content of uranium thorium and potassium in the Earth.

Table 1

Isotope	Period of Semi-Decomposition, years	Initial Heat Emission, cal/g · year	Isotope	Period of Semi-Decomposition, years	Initial Heat Emission, cal/g · year
Be^{10}	$2.5 \cdot 10^6$	$5.3 \cdot 10^{-5}$	J^{129}	$1.7 \cdot 10^7$	$3.9 \cdot 10^{-7}$
Al^{26}	$7.2 \cdot 10^5$	$5.6 \cdot 10^{-2}$	Np^{237}	$2.2 \cdot 10^6$	$7.2 \cdot 10^{-5}$
Cl^{36}	$3 \cdot 10^5$	$6.6 \cdot 10^{-3}$	Pu^{244}	$7.6 \cdot 10^7$	$5.1 \cdot 10^{-7}$
Fe^{60}	$\sim 3 \cdot 10^5$	$\sim 0 \cdot 12$	Cm^{247}	$> 4 \cdot 10^7$	$<2.1 \cdot 10^{-7}$

a) The short-lived isotopes. Presently, 27 short-lived /98 isotopes are known primarily through the discovery of their products of decomposition in abnormally large quantities in comparison with the curve of their cosmic abundance. Many short-lived isotopes are obtained artificially, and the periods of their decomposition are known exactly. They constitute 10^6 - 10^8 years which is significantly less than the age of the Earth. The emission of energy during the decomposition of the short-lived isotopes could have played a substantial role in the early period of the Earth's history.

Until recently all isotopes with periods of 10^8 years and less have decomposed to immeasurably small quantities. This is the way in which the introduction of the term "extinct radioactivity" is explained.

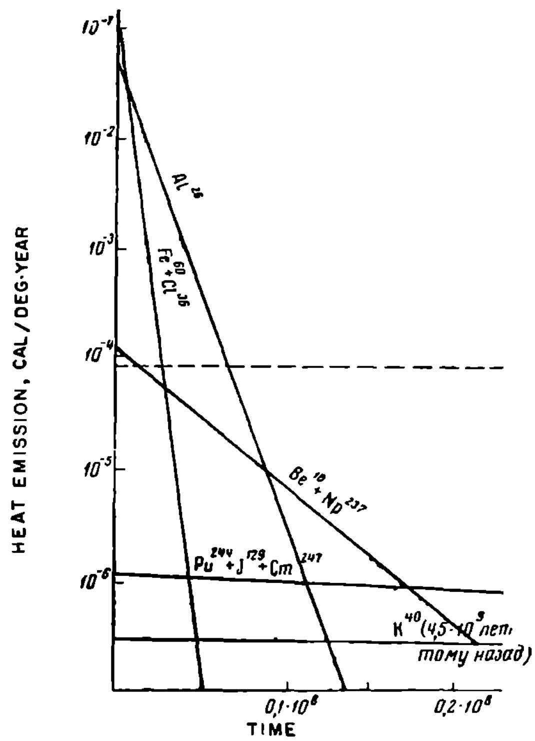


Fig. 1. The Emission of Heat by the Short-Lived Isotopes.

In Table 1, those short-lived isotopes are listed which have the longest life time and show the greatest energetic effect. In Fig. 1 the

emission of heat with time from the short-lived isotopes is shown. The beginning of time is taken from the moment $4.5 \cdot 10^9$ years ago. We see that, at first, the greatest emission of heat was caused by isotopes Cl^{36} and Fe^{60}. But their activity continued only for a million years after the nuclear genesis. After this, during the course of 10 million years, the greatest heat emission was generated by Al^{26}. During the following 10 million years, Be^{10} and Np^{237} were dominant. The data for the construction of Table 1 and Fig. 1 were borrowed from the article of Fish, Goles, and Anders (Ref. 1). Since the periods of the semi-decomposition of the more effective short-lived isotopes do not exceed 10^6 - 10^7 years, then their contribution to the generation of heat could be accomplished if the moment of the beginning of the Earth's formation were very close to the moment of nuclear genesis. In such a case, the contribution of short-lived isotopes would be maximal provided that the time of planetary formation were less than or equal to the lifetime of the isotopes. However, according to the present estimates, it is much greater. /99

The period of planetary formation is dependent upon the rate of their growth. Concrete schemes of the physical process of the accumulation were proposed, and the curves of the rate of planetary growth were revealed by Safronov (Refs. 2, 3). The curve of the Earth's growth is presented in Fig. 2. Along the axis of the ordinates is the magnitude $(m/M)^{1/3}$ where m is the mass of the growing Earth, and M is the present mass. We see that the growth of the Earth was mainly concluded in the course of the first 10^8 years. Thus, m = 0.99 M is reached where t = $2.5 \cdot 10^8$ years.

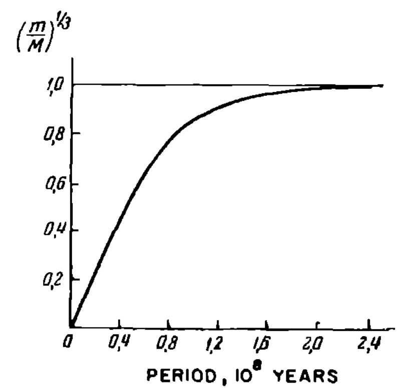


Fig. 2. The Rate of the Growth of the Earth.

Let us compare the rate of the Earth's growth with the period of activity of the short-lived isotopes. The period of their activity does not exceed the interval $0.2 \cdot 10^8$ years. According to the rate of growth of the Earth's mass, the magnitude $(m/M)^{1/3}$ up to the moment of time $t = 0.1 \cdot 10^8$ or $0.2 \cdot 10^8$ years should be equal to $(m/M)^{1/3} = 0.1$ or 0.2, respectively. From this we get $m = 0.001M$ or $0.01M$. The latter value gives approximately the magnitude of the entire mass of the Earth's inner core. In this way, the emission of heat by the short-lived isotopes can be coordinated only with the small initial nucleus of the Earth which occupies only a small volume of the Earth's inner core.

From Fig. 1, we see that in the interval of $0 - 0.1 \cdot 10^8$ years, the main portion of the generation of heat was determined by isotopes Al^{26}, Cl^{36}, and Fe^{60}. They could guarantee the heating of substance to the melting point (the dashed line in Fig. 1 shows the energy necessary for the increase of temperature to 3000°C). At this time, the radius of the Earth's nucleus amounted to less than 100 kilometers and the nucleus was able to be melted. But the heat could not hold its ground within such a small body and was quickly lost in space (Ref. 4). Therefore, one can disregard its influence on the upper layers.

After the lapse of $0.1 \cdot 10^8$ years, the deposit in the generation of heat from Al^{26} became insignificantly small; the main part of the generation was for Be^{10} and Np^{237}. The maximal temperature which could be reached within the time $t - t_1$ is determined by the formula

$$T = \frac{1}{c} \int_{t_1}^{t} \sum_i H_i e^{-\lambda_i \tau} \, d\tau ,$$

where λ_i is the constant of decomposition of the i^{th} element and H_i /100 is the initial content of it. In the given case

$$T = \frac{1}{c} \frac{H_{Be}}{\lambda_{Be}} \left(e^{-\lambda_{Be} t_1} - e^{-\lambda_{Be} t}\right) + \frac{1}{c} \frac{H_{Np}}{\lambda_{Np}} \left(e^{-\lambda_{Np} t_1} - e^{-\lambda_{Np} t}\right).$$

By substituting $t_1 = 0.1 \cdot 10^8$; $t_2 = 0.2 \cdot 10^8$ years; $H_{Np} = 7.2 \cdot 10^{-5}$ cal/g · years; $\lambda_{Np} = 0.314 \cdot 10^{-6}$ years^{-1}; $H_{Be} = 5.3 \cdot 10^{-5}$ cal/g · years; $\lambda_{Be} = 0.276 \cdot 10^{-6}$ years^{-1}; $c = 0.2$ cal/g · deg. We find that $T \approx 100$°C.

In that way, provided that the moment of the formation of the elements coincides with the beginning of the Earth's formation, the maximal effect of the short-lived isotopes should have been reduced to the melting of the center part of the Earth's nucleus which makes up one tenth of the mass of the inner core. Also this must have been reduced to a rapid drop in temperature to the boundary between the internal and external nucleus.

During the formation of the region of the external nucleus, the radioactivity of the short-lived isotopes had already practically vanished. In this way, if the effect of these isotopes were told in the early history of the Earth, then the melted area should be the inner nucleus rather than the outer. However, the geophysical facts indicate the fact that the inner nucleus of the Earth should now be firm and the outer melted.

Just like Fish, Goles, and Anders (Ref. 1) we come to the deduction that the action of the short-lived isotopes can be appreciable only for very small planetary nuclei-"planetesimals". Their cores can pass through the state of melting. These planetesimals can then fall out on to the growing nucleus of the Earth bringing on themselves traces of melting, but the entire Earth on the whole could not have been melted.

b) The emission of gravitational energy; the release of elastic energy. Certain authors (Refs. 5 through 7) consider that the emission of gravitational energy during the process of planetary formation can be the source of the initially molten condition of the Earth. The potential gravitational energy which is being released during the condensation of an object of mass M and radius R is equal to

$$W = -\frac{3}{5}\frac{GM^2}{R}\,. \tag{1}$$

Latimer (Ref. 6) indicated that during the accretion of the Earth from a cloud of dust energy must have been released on the order of $4 \cdot 10^4$ J/gm of matter. Fesenkov (Ref. 5) has calculated according to formula (1) that W for a homogeneous sphere with a mass and radius of the Earth is equal to $W \approx 4 \cdot 10^{39}$ erg. The more careful estimate of Urey (Ref. 8) and Beck (Ref. 9) conducted with the calculation of the uniform distribution of density gives $W \approx 2.5 \cdot 10^{39}$ erg. Verhoogen (Ref. 7) indicates that this figure is equivalent to the emission of 9000 cal/gm of heat. /101
This is sufficient for the complete melting of the entire Earth which, as he considers, has taken place. Fesenkof on the basis of the virial theorem considers that half of the energy should have been converted into

heat. It is necessary to note, however, that this theorem is inapplicable in the case of the precipitation of bodies on a hard surface.

The large magnitude of the potential gravitational energy still cannot serve as any kind of basis for the initially melted condition of the Earth. Actually, during the union of the particles, this energy is partially converted into the internal energy of the substance increasing its temperature and deformation and partially being lost through radiation. The portion which should go into the increase of the temperature is substantially dependent on the speed of accumulation and the speed with which heat can be radiated in space.

The examination of concrete schemes of the process of accumulation led V. S. Safronov (Refs. 2, 3, 10) to the deduction that the enormous energy which the particles, that have been striking against the surface bear with them, must be rapidly reradiated in space. Only an insignificant portion of the energy went for heating the Earth.

Later, we shall show that a significant portion of the energy W should have been expended also for the execution of only one elastic deformation of the self-gravitating globe of the Earth. Actually, the elastic energy $\Delta\epsilon$ of an element with a volume of ΔV of a hard body characterized by the elastic moduli K (the modulus of compression) and μ (the modulus of rigidity) is given by the formula (Landau and Lifshitz, Ref. 11)

$$\Delta\epsilon = \frac{K}{2} U_{\ell\ell}^2 \Delta V + \mu \left(U_{ik} - \frac{1}{3} U_{\ell\ell} \delta_{ik} \right)^2 \Delta V, \tag{2}$$

where U_{ik} is the sum of the components of the tensor deformation; $U_{\ell\ell}^2$ is the sum of the squares of the components of the tensor and $\delta_{ii} = 3$. In the case of spherical symmetry $U_{ik} = 0$ $(i \neq k)$

$$U_{\ell\ell}^2 = \left(\frac{\partial u}{\partial r} \right)^2 + 2 \left(\frac{u}{r} \right)^2 , \tag{3}$$

where u is the shift in the radial direction. The full inner elastic energy of a homogeneous ball with a radius R is determined by the integration of the expression (2) according to the volume:

$$\epsilon = \iiint_V \left(\frac{K}{2} + \mu \right) U_{\ell\ell}^2 \, dV = 4\pi \int_0^R \left(\frac{K}{2} + \mu \right) \left[\left(\frac{\partial u}{\partial r} \right)^2 + 2 \left(\frac{u}{r} \right)^2 \right] r^2 dr. \tag{4}$$

The shift $\vec{u}$ in a gravitating elastic medium is determined as a solu- /102
tion of the equation of equilibrium of the elastic medium

$$\text{grad div } \vec{u} - \frac{1 - 2\sigma}{2(1 - \sigma)} \text{ rot rot } \vec{u} = - \rho\vec{g} \frac{(1 + \sigma)(1 - 2\sigma)}{E(1 - \sigma)}, \quad (5)$$

where E is Young's modulus and σ is the Poisson coefficient. By a spherical symmetry for a homogeneous sphere, we have

$$\text{grad div } \vec{u} = \frac{d}{dr}\left[\frac{1}{r^2}\frac{d(r^2\vec{u})}{dr}\right],$$

$$\text{rot } \vec{u} = 0,$$

$$\vec{g} = - g\frac{\vec{r}}{R}.$$

Then (5) will be reduced to the equation

$$\frac{E(1 - \sigma)}{(1 + \sigma)(1 - 2\sigma)} \frac{d}{dr}\left[\frac{1}{r^2}\frac{d(r^2 u)}{dr}\right] = \rho g \frac{r}{R}. \quad (6)$$

A marginal condition will be the absence of normal strains on the surface $r = R$:

$$\sigma_{rr}\Big|_{r=R} = \frac{E}{(1 + \sigma)(1 - 2\sigma)}\left[(1 - \sigma)\frac{\partial u}{\partial r} + 2\sigma\frac{u}{r}\right] = 0. \quad (7)$$

The final solution when $r = 0$ has the form

$$u = - g\rho R \cdot \frac{(1 - 2\sigma)(1 + \sigma)}{10E(1 - \sigma)} r \left[\frac{3 - \sigma}{(1 + \sigma)} - \frac{r^2}{R^2}\right]. \quad (8)$$

We estimate the energy ϵ when $\sigma = 0.25$:

$$\epsilon = 4\pi \int_0^R \left(\frac{K}{2} + \mu \right) \left\{ \left[\frac{g\rho R}{E} \left(0.25 \frac{r^2}{R^2} - 0.183 \right) \right]^2 + \right.$$

$$\left. + 2 \left[- \frac{0.625}{7.5} \frac{g\rho R}{E} r \left(2.2 - \frac{r^2}{R^2} \right) \right]^2 \right\} r^2 dr. \qquad (9)$$

Assuming that $K = 2 \cdot 10^{12}$ dynes/cm^2, $\mu = 1 \cdot 10^{12}$ dynes/cm^2, $E = 1.5$ K, $g = 1000$ cm/sec^2, $\rho = 5.5$ gm/cm^3 we obtain

$$\epsilon = 0.5 \cdot 10^{39} \text{ erg.} \qquad (10)$$

This energy according to its magnitude composes 1/5 of the full gravitational energy of the Earth's sphere. Relative to what we have said, the complete melting of the Earth due to the release of gravi- /103 tational energy is only slightly probable. It is difficult to release the potential energy used up in the execution of the deformation under gravitational action. It could have been released only if the complete relaxation of the elastic strains had occurred. An insignificant relaxation, of course, could have taken place. This is undoubtedly attributed to the tangential strains but to a certain degree to the normal strains also. In the process of the relaxation, heat should have been emitted. But this process was slow. One can pose the question concerning the new source of heat in the interior of the Earth which arises as a result of the release little by little of the elastic energy in the process of the elastic strains. In the article "The Thermo-Elastic Strains Within the Earth's Sphere (Ref. 12) we proved that the release of only the energy of thermo-elastic deformation arising due to the uneven distribution of temperature in the Earth can guarantee the annual emission of seismic energy during the existence of the Earth. However, the quantitative solution of the problem concerning this source of heat is complicated by the absence of experimental data concerning the times of relaxation in the Earth.

c) The emission of energy in gravitational differentiation. Ye. N. Lyustikh (Ref. 13), Urey (Ref. 8) and V. A. Krat (Ref. 14) calculated that during the changeover of the Earth from a more or less homogeneous silicate-metallic mixture into the contemporary Earth with a nucleus of iron which has a more stable state (the state of equilibrium) a certain portion of gravitational energy must have been emitted. This portion has turned out to be equal to the type in the following equation

$$\Delta W = \begin{cases} 1.5 \cdot 10^{38} \text{ erg (Lyustikh)} \\ 2 \cdot 10^{38} \text{ erg (Urey)} \\ 5 \cdot 10^{38} \text{ erg (Krat)} \end{cases}$$

Depending on the figure $2 \cdot 10^{38}$ erg, Ringwood (Ref. 15) concluded that the emission of energy during the gravitational differentiation was enough for the passage of the Earth's sphere through the molten state. Ringwood proposes that the process of the metal's introduction into the nucleus was catastrophic and was executed by means of a convective cataclysm. Lyustikh and Krat consider that the emission of the energy occurred gradually during the life of the Earth.

A simple calculation shows that this energy was insufficient for the complete melting of the Earth even if it were instantaneously emitted. The gravitational energy which was uniformly emitted throughout the entire Earth would have called for an increase in temperature of the Earth's sphere on a magnitude of

$$\frac{\Delta W}{M \cdot c} = \frac{(1.5 \div 2.0) \cdot 10^{38} \text{ erg}}{6 \cdot 10^{27} \cdot 0.3} = 1900 \div 2600^\circ,$$

which is close to the melting temperature in the upper layers of the /104
Earth but significantly lower than the melting temperature of the substance in the Earth's crust. If you consider that the energy was emitted only in the nucleus of the Earth, then the temperature of the nucleus could have been raised on a magnitude of

$$\frac{\Delta W}{m_n c} = \frac{(1.5 \div 2.0) \cdot 10^{38} \text{ erg}}{1.9 \cdot 10^{27} \cdot 0.3} = 6000 \div 8000^\circ.$$

This could have been entirely adequate for the melting of the Earth's nucleus. Urey (Ref. 8) has already indicated this. However, there is no suitable explanation as to why the energy of the gravitational differentiation occurring throughout the entire Earth should have been emitted only in the region of the nucleus. The emission of energy, also, should have occurred throughout the entire volume. If it is clear that the nucleus of the Earth consists of iron, then the change of gravitational energy during the shifting of the Earth's masses into a position more suitable from the viewpoint of equilibrium (as is the change of this energy during the decrease or increase of the radius of the Earth under the effect of temperature) should be taken into consideration in the

examination of the thermal evolution of the Earth. Then it is necessary to take into consideration that this process was not instantaneous and, in a rough approximation, one can consider that the emission of heat occurred as if with a uniformly distributed source with an intensity $q = (\Delta W/mt)$ cal/gm-sec, where t is the period of the existence of the process of gravitational differentiation, and m is the mass of that volume where the process occurs.

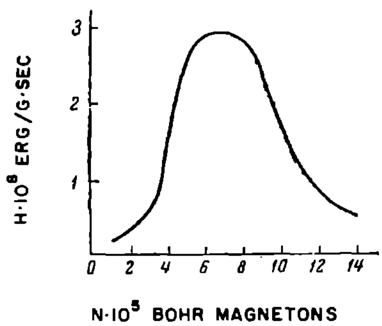


Fig. 3. The Intensity of the Emission of Energy Dependent on the Magnitude of the Magnetic Moment of the Neutrino.

d) The hypothesis of the neutrino. Ya. B. Zeldovich and Ya. A. Smorodinskiy (Ref. 16) indicated that the energy density of an unobserved form (density of neutrino, graviton, etc.) can be not less than the density of common matter 10^{-29} gm/cm^3 (the average density of the nucleons in the universe).

According to Bethe (Ref. 17) in the nitrogen-carbon cycle process of the Sun, neutrino particles should be emitted during the decomposition of N^{13} and O^{15}. Seven percent of all energy lost by the Sun should occur from neutrinos. The flux of neutrinos from the Sun should have a value of $7.4 \cdot 10^{10}$ neutrinos/cm^2 · sec. Cormak (Ref. 18) stated the assumption that if the neutrinos were to possess magnetic moments, they should have interaction with the electrons of the metallic iron nucleus of the Earth

through a kind of inelastic collision. Thereby, he evaluated the intensity of the energy emission as it is dependent upon the magnitude /105 of the magnetic moments of the neutrinos (Fig. 3). Certain investigators hold that the upper limit for n has a magnitude of $n_0 = 180 \cdot 10^{-6}$ Bohr magnetons. In that way, the probability of the peak (on the curve - Fig. 3) value of energy in $3 \cdot 10^{-8}$ erg/gm · sec = $0.7 \cdot 10^{-15}$ cal/gm · sec is not excluded. This energy, which corresponds to the magnetic moment and $n = 8 \cdot 10^{-6}$ Bohr magnetons is not excluded. The heating, due to the absorption of neutrinos during the existence of the Earth can, in this case, give

$$T = \frac{Q \cdot t}{c} = \frac{0.7 \cdot 10^{-15} \cdot 4.5 \cdot 10^{9} \cdot 3.15 \cdot 10^{7}}{0.3} \approx 300^{\circ},$$

i.e., a practically small value.

e) The emission of heat due to tidal friction. The simple calculation made by P. P. Cotov (Ref. 19) showed that the change in the duration of a day within the lifetime of the Earth from 4 hours to 24 hours leads to the decrease of energy of rotation from $107 \cdot 10^{36}$ erg to $3 \cdot 10^{36}$ erg. The difference in the energies is $104 \cdot 10^{36}$ erg. During the departure of the Moon from the Earth, $13.5 \cdot 10^{36}$ erg should have been depleted. The remaining $9 \cdot 10^{37}$ erg should have been in some way absorbed inside the Earth. This energy has a ratio of 5:1 less than the energy of radiogenic origin. This is the energy which is emitted during the lifetime of the Earth. Therefore, the contribution of this effect in the energetic balance of the entire Earth's sphere is insignificant. Ye. N. Lyustikh (Ref. 20) and V. A. Magnitskiy (Ref. 21) have indicated that the deceleration of the Earth's rotation is mainly due to the tidal friction in the seas. But a small part of this energy can be absorbed by the hard part of the Earth owing to its deviation from ideal elasticity. This energy undoubtedly is less than the energy of gravitational differentiation.

The long-lived isotopes. In recent years much work has been dedicated to the detailed investigation of the role of U^{238}, AcU^{235}, Ta^{232}, and K^{40} in the thermal evolution of the Earth (Refs. 12, 15, 22-32). Until now, the content of radioactive elements in the Earth's sphere appears to be the greatest undetermined parameter in this problem. For its determination the mean composition of the Earth is usually identified with the mean composition of the meteorites. Actually Birch (Ref. 31) and MacDonald (Refs. 28, 32), whose works are based on the

semi-emperical equation of state for matter in the Earth's crust, come to the conclusion that by the contents and elastic properties, the composition of the crust most closely approximates the substance of meteorites of the chondrite type. Many authors have usually identified the nucleus of the Earth by its composition with iron meteorites (Refs. 7, 22). The latest experiments on the expansion of shock waves in rock formations have proved to be in greater agreement with the supposition concerning the iron nucleus and not silicate which is transferred into a metallic phase (Refs. 33, 34).

In our first calculations of the thermic history of the Earth (Ref. 25), the schemes of the silicate nucleus were investigated more frequently than the schemes of the iron nucleus. Two extreme /106
evaluations A and B have been distinguished for the mean concentration of radioactive elements throughout the entire Earth. This concentration in the initial stages of the development of the Earth was considered to be uniform. Then the ratio of iron to rock meteorites was taken as 1 to 3, which corresponds to the iron content in the volume of the Earth's nucleus. As a result (Ref. 27) the value C for which this ratio equaled 1 to 7 was utilized, which corresponded more nearly to the silicate composition of the nucleus and to the portion of iron which was scattered throughout the volume of the Earth in the form of oxides.

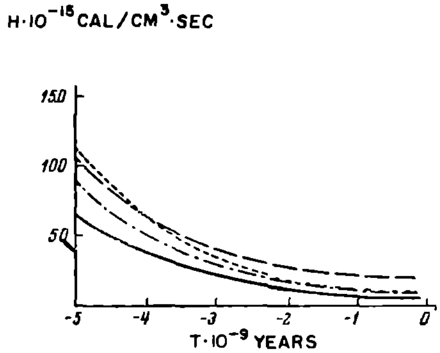


Fig. 4. The Curves of the Generation of heat by long-lived isotopes. The Continuous line is the value A'. The dotted line is A, the broken line is C and the dash and dot line is C/2.

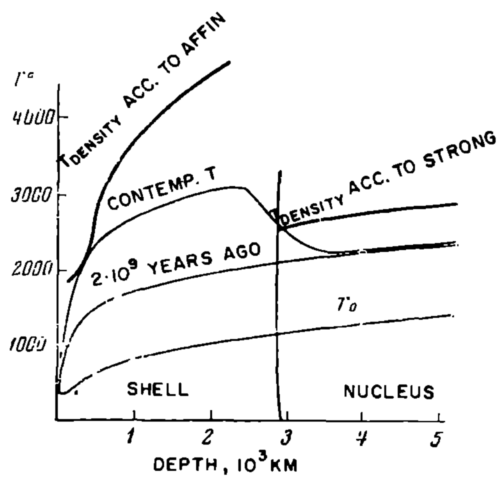


Fig. 5. The Scheme of the Distribution of Temperature in the Earth with an Iron Nucleus According to the Calculation of the Energy of Differentiation (in $^{\circ}$K).

According to data from the new method of discovering traces of radioactivity by means of the radiation activation analysis, it is established that in the iron meteors there should be a significantly smaller content of U^{238} and Th^{232} than was considered earlier. From these data, it also follows that among our evaluations, A appears to be much more probable, but it should be made more exact in relation with the data of the radiation activation analysis on iron meteorites.

The content of U and Th in rock meteorites according to the evaluation of A was taken by us to be equal to $1.3 \cdot 10^{-8}$ gm/gm U, $3.9 \cdot 10^{-8}$ gm/gm Th. This was practically confirmed by recent data (Ref. 35) which give $1.1 \cdot 10^{-8}$ gm/gm U and $4.4 \cdot 10^{-8}$ gm/gm Th. By defining the new evaluation by A' we shall have $0.87 \cdot 10^{-8}$ gm/gm U; $3.5 \cdot 10^{-8}$ gm/gm Th and $8 \cdot 10^{-4}$ gm/gm K.

The corresponding generation of heat is shown in Fig. 4. As a comparison, the curves of the generation according to our previous values AC and also the evaluation of C/2 utilized by B. Yu. Levin and S. V. Mayeva (Ref. 30) were offered. It is obvious for the Earth with an /107 iron nucleus, the curve of the heat emission should be below the results obtained earlier. For a heat due to radioactivity, the heat accumulation should be less, and the resulting temperatures should be lower. The temperature variation was evaluated for the model of the Earth containing an iron core. However, in addition to radioactivity the heat emitted in the process of gravitational differentiation should have been considered in the calculations.

On the basis of the recent paleomagnetic reduction (Ref. 36) the magnetic field of the Earth, approximately of the same force as it is now, should have already existed $2 \cdot 10^9$ years ago. This means that the formation of the nucleus must have been concluded before this period. If we accept, as before, the age of the Earth being equal to $4.5 \cdot 10^9$ years we consider that the changeover of the Earth with the formation of an iron nucleus occurred during the first $2.5 \cdot 10^9$ years. During this time concentration of the radioactive elements in the nucleus of the Earth was reduced to zero as follows from data given in the paper (Ref. 35). During the last $2 \cdot 10^9$ years the temperature of the nucleus could have risen only with the influx of heat across its boundaries from internal layers of the crust. In Fig. 5 the temperature curves corresponding to the described scheme are shown. They give the minimum temperatures due to radioactivity. If we take into consideration the fact, that in the process of formation of the nucleus during the first 2.5 billion years, the emission of heat

in the process of gravitational differentiation of equal intensity, has been added to the radioactive heat generation (equally throughout the Earth), then T should be increased to 2000°C. We see that as a result of the influx of heat from the crust, the temperature of the external parts of the nucleus in this new scheme could reach the melting point. Thus, it is not obligatory for the explanation of the external liquid nucleus to resort to the hypothesis of the initially molten and slowly cooling Earth as did Jacobs (Ref. 37) and Verhoogen (Ref. 7) recently.

From what we have said, it follows that the evolution of the Earth was directed toward a gradual secular heating under the effect of heat of slowly disintegrating radioactive elements and possibly of the gravitational differentiation of the Earth.

REFERENCES

1. R. A. Fish, G. G. Goles, E. Anders. Astrophysical Journal 132, 243, 1960.

2. V. S. Safronov. Astronomicheskiy Zhurnal 31, 499, 1954.

3. V. S. Safronov. Voprosy Kosmogonii (Problems of Cosmogony) 6, 63, 1958.

4. Ye. A. Lyubimova, A. G. Starkova. Astron. Zh. 31, No. 5, 429, 1954.

5. V. G. Fesenkov. Astron. Zh. 34, No. 1, 105-120, 1957.

6. W. M. Latimer. Science 112, 101-104, 1950.

7. J. Verhoogen. Science 25, No. 1, 1960; Geophysical Journal of the Royal Astronomical Society 4, 71, 1961.

8. H. Z. Urey. The Planets, Their Origin and Development, London 1952.

9. A. E. Beck. Journal of Geophysical Research 66, No. 5, 251, 1961.

10. V. S. Safronov. Izvestiya Akademii Nauk, SSSR, Seriya Geofizicheskaya, No. 1, 139, 1959. /108

11. L. D. Landau, Ye. M. Lifshits. Mekhanika Sploshnikh Sred (The Mechanics of Continuous Matter), Moscow, Gosudarstvennoye Izdatel'stvo Tekhnicheskoy i Teoreticheskoy Literatury, 1953, p. 629.

12. Ye. A. Lyubimova. Trudy Instituta Fiziki Zemly, No. 20 (187), 1961.

13. Ye. N. Lyustikh. Doklady Akademii Nauk, SSSR, 69, No. 8, 1417, 1948.

14. V. A. Krat. Voprosy Kosmogonii (Problems of Cosmogony) 7, 97, 1960.

15. A. E. Ringwood. Geochim. et Cosmochim. Acta 20, No. 3-4, 241, 1960.

16. Ya. B. Zeldovich, Ya A. Smorodinskiy. Zhurnal Experimentalnoy i Teoreticheskoy Fiziki 41, No. 3 (9), 1961.

17. K. Bethe. Phys. Rev. 55, 434, 1939.

18. N. Cormak. Phys. Rev. 95, No. 2, 1954.

REFERENCES (Cont'd)

19. P. P. Botov. Priroda No. 3, 122-124, 1960.

20. Ye. N. Lyustikh. Priroda No. 3, 123, 1960.

21. V. A. Magnitsky. Priroda No. 3, 123, 1960.

22. J. A. Jacobs. Travaux, Sci. Ser. A, France 18, 1954.

23. J. A. Jacobs. Nature 185, No. 4712, 521, 1960.

24. J. A. Jacobs and D. W. Allan. Transactions, Royal Society of Canada, 48, Ser. III, 1, 33, 1954.

25. Ye. A. Lyubimova. Izvestiya Akademii Nauk SSSR. Ser. Geofiz. No. 2, 3, 1952.

26. Ye. A. Lyubimova. Doklady A. N. SSSR, 107, No. 1, 55, 1956.

27. H. A. Lubimova. Geophys. J. R. A. S. 1, No. 2, 1958.

28. G. J. F. Mac Donald. J. Geophys. Res., 61, No. 2, 387, 1956.

29. S. P. Clark. J. Geophys. Res., 66, No. 4, 1231, 1961.

30. B. Yu. Levin, C. V. Mayev. Izv. A. N. SSSR. Ser. Geofiz. No. 2, 243, 1960.

31. F. Birch. Bulletin of the Geological Society of America 69, 483-486, 1958.

32. G. J. F. Mac Donald. J. Geophys. Res. 64, No. 11, 1959.

33. L. V. Altshuler, K. K. Krupnikov, B. N. Lebedyev, S. B. Kormer, and M. I. Brazhnik. Zh. Eksperim. i Teor. Fiz., 7, 606, 1958.

34. L. V. Altshuler, S. B. Kormer. Izvestiya A. N. SSSR, Ser. Geofiz. No. 1, 33, 1961.

35. H. Hamaguchi, G. W. Reed and A. Turkevich. Geochim. et Cosmochim. Acta 12, 337-347, 1957.

36. A. Coxa, R. R. Doell. Bull. Geol. Soc. Am. 71, 645-768, 1960.

37. J. A. Jacobs. Geophys. J. R. A. S. 4, 84, 1961.

REVIEW OF CONTEMPORARY DATA CONCERNING THE MOON /109

By

B. Yu. Levin and Ye. L. Ruskol

This review includes the contemporary data concerning the motion of the Moon, its atmosphere, photometric characteristics, surface temperature, nature of the surface layer, relief, origin of the Moon, its thermal history and internal structure.

I. The Orbital Motion, Rotation and Librations

1. The motion of the Moon about the Earth which takes place in the first approximation, along an elliptical orbit, is extremely complicated. This complication is explained by the enormous disturbances from the Sun, a large part of which is connected first with the proximity of the Earth (and the Moon) to the Sun and second with the comparatively great remoteness of the Moon from the Earth (in comparison with satellites of other planets). At the same time, the proximity of the Moon to the Earth (in comparison with other celestial bodies) permits a very accurate observation of its motion, and in an accurate theory this compels the consideration of even small disturbances from Venus, Mars and Jupiter and also disturbances coming from the oblate shape of the Earth.

2. Due to the fact that the period of axial rotation of the Moon is equal to the period of its revolution about the Earth, approximately half of the lunar surface is observable from the Earth. However, the inclination of the lunar equator to the plane of its orbit creats an optical libration along the latitude permitting a view of somewhat more of the Moon in the regions of its poles. The uniform rotation of the Moon in conjunction with the varying velocity of its motion in orbit from the law of areas (the second law of Kepler) creates an optical libration along the longitude enabling us to see more of the Moon in its regions along the east and west boundaries. Because of the powerful /110 perturbations of the Moon's orbit, the magnitude of the librations does not remain constant but changes somewhat. The maximum libration in latitude is $6^\circ 50'$ and the maximum libration in longitude is $7^\circ 54'$. Thus the general optical libration can exceed 10° in rare cases.

Furthermore, there is also a parallactic libration. By observing the Moon not from the center of the Earth but from its surface, and also by observing it not while it is at the zenith but at the horizon, we can, in addition, see more because of the daily parallax (about 1°).

As a result of these reasons, one can observe 59% of the surface of the Moon, while only 41% of its surface is constantly turned toward the Earth.

Because of the libration, the center of the visible disc of the Moon shifts along its surface. The speed of this shift is determined primarily by the shift of the observer. This shift is relative, due to the diurnal rotation of the Earth, i.e., by the change of parallactic libration. The diurnal changes are such that the speed of the shift is at its maximum during the culmination of the Moon and is at its minimum during its rising or setting. (In addition, this speed is essentially dependent on the latitudes where the observer is located and the inclination of the Moon.)

3. The equality of the periods of rotation and revolution of the Moon is not an accidental coincidence but a stable state stipulated by the triaxial nature of the ellipsoid of lunar inertia. This stable state was reached as a result of a gradual braking of the once rapid axial rotation of the Moon by mundane tidal forces. Thereby, in accordance with stability requirements, a small axis of the inertia ellipsoid proved to be directed on Earth, i.e., axis along which the contour of the Moon is most elongated.

4. The optical libration (with the exception of a small part connected with the daily parallax) leads to the fact that the major axis of the lunar contour (i.e., the small axis of the inertia ellipsoid) declines at times from the direction toward the Earth and then there appears a pair of forces striving to return this axis to the radius vector. As a result, a constrained physical libration occurs, i.e., an actual lunar oscillation of an extremely small amplitude (approximately 2'). Nevertheless a study of physical libration is of great interest, since it enables us to identify the difference of the equatorial axes of the Moon's ellipsoid of inertia. Generally speaking, a free libration of the Moon is also possible, but it is so small that it cannot be observed.

5. The exact determination of the lunar contour from optical observations is very difficult. The contour of the Moon is distorted by irregularities at the rim of the lunar disc, which vary in the presence of different librations, and the determination of the most probable /111
curvature which represents the lunar disc, requires prolonged systematic observations. A. A. Yakovkin (Refs. 53, 54) came to the conclusion that the contour of the Moon is asymmetrical. The visible contour of the

northern hemisphere approaches a circle, whereas the contour of the southern hemisphere is better described by an ellipse with the major semiaxis oriented at a small angle toward the southern polar axis. The deviation of this ellipse from a circle (ellipticity) is approximately 1:1300. However, the observations of the lunar profile by Vaymer (Ref. 5) showed that the smoothed contour of the lunar disc can be presented as a circle, and that the conclusion concerning the asymmetry of the Moon's contour has no basis. An indication of another type of asymmetry: a weak ellipticity (1:1200) with the major axes inclinated at an angle of 36° toward the polar axis of the Moon, was obtained from a study of a single observation of the Moon during its full stage [Potter (Ref. 31)].

Evidently the deviation of the Moon's contour from a circle is specified by its relief and is not related to its dynamic shape.

Baldwin (Ref. 57) utilizing determinations by Frantz and Saunder of the heights of points on the lunar surface and while investigating separate points located on the continents and seas, found that the average level of the seas is by 1.5 km lower than the level of the continents and that the protuberance directed toward the Earth has an altitude of 2 km. However, this result is very inaccurate.

As a result of low accuracy of previous measurements on old hypsometric charts of the Moon [Frantz (Ref. 69), Ritter (Ref. 103), Schrutka-Rechtenstamm and Hopmann (Ref. 105)], the isohyps traverse the seas and continents, without exposing the seas in the capacity of depressions and the continents in the capacity of protuberances [see also Du-Fresne, (Ref. 65)]. This appears only on the new hypsometric chart of the Moon which was recently published by Baldwin (Ref. 58) and which is based on his personal measurements of the elevations. However, the elevations as such and the general data obtained from them concerning the lunar contour have not as yet been published.

6. It has been proven for a long time that the dynamic contour of the Moon, i.e., its ellipsoid of inertia does not correspond to the hydrostatic equilibrium at the present character of its motion. The theory of the Moon's motion, makes it possible for us to find, with great accuracy, the relation $\frac{C - A}{C}$ (A, B, C are the moments of inertia in relation to the main axes), according to the inclination of the lunar equator toward the elliptic.

$$\frac{C - A}{C} = 0.00063,$$

which is 17 times greater than the equilibrium value (0.0000375). The observations of the amplitude of the Moon's physical vibration permits us

to find the relationship $\frac{B - A}{C}$ but with little accuracy. We obtain/112

$$\frac{B - A}{C} = 0.00012 \div 0.00020$$

instead of the equilibrium value 0.0000281. The difference in the equatorial axis of inertia is several times less than the difference of the polar and equatorial axis; as in the case of the tidal raise, these differences should be almost identical at any distance between the Moon and the Earth, since during the hydrostatic equilibrium and the equality of the periods as well as rotations are always $\frac{C - A}{B - A} = \frac{4}{3}$. The basic deviation of the dynamic contour of the Moon from the equilibrium is its oblate nature on which only a small elongation along the direction to the Earth is superimposed.

If the Moon had a homogeneous density, then the difference of its axis of inertia would signify that the polar diameter is approximately 2 km less than the equatorial diameter and that the difference in the equatorial diameters is about 0.3 km. Thereby there should exist in the center of the Moon, stresses of about 20 km/cm^2, which is only possible in the presence of a lunar rigid state of matter. However, according to present concepts, the central parts of the Moon should be heated, and not molten, then in any case they would be softened and therefore incapable of withstanding similar stresses over a period of several million years. Therefore Urey, Elzasser, and Rochester (Ref. 114) have considered a model of the Moon, in which the stresses at the center equalled zero and the observed triaxial nature of the ellipsoid of inertia is caused by a varying distribution of density along various radii. (The density is at its maximum along the polar axis and at its minimum in the direction towards the Earth.) Thus, the absence of stresses in the center is achieved in this model by means of the introduction of such stresses into the shell, for which it is easier to admit a solid state.

In the future, studies on the mass distribution of the Moon will be accomplished by means of observations from aboard artificial lunar satellites.

II. The Atmosphere

7. Because of the Moon's small mass, the force of gravity on its surface is 6 times less than on the Earth ($g_{Moon} = 0.165\ g_{Earth} = 162$ cm/sec^2). Also of a correspondingly lower magnitude is the velocity of escape, or according to rocket terminology the second cosmic velocity

(v_∞ = 2.38 km/sec). This velocity is less than 5 times greater than the mean quadratic velocity of the molecules of common atmospheric gases and hence the condition necessary for the existence of a sufficiently /113 stable atmosphere is not fulfilled. Therefore, the Moon is practically deprived of an atmosphere.

8. The absence of any perceptible atmosphere on the Moon is apparent in the blackness of shadows and the sharpness of all details. This is also evident in the absence of the phenomena of refraction in the occultation of stars by the Moon, this is most apparent during observations of the points at the narrow sickle of the Moon (about new Moon). At this time the conditions of illumination are such that even a rarefied atmosphere should have given a perceptible scattered light. The absence of this light showed that the density of the atmosphere near the surface of the Moon is at least 10^9 less than the density of the atmosphere at the surface of the Earth [Dollfus (Ref. 63)].

9. If C, N, O, H were present in the lunar atmosphere in a sufficient quantity, then the florescence of neutral and ionized molecules, primarily CO^+ and N_2^+ might be observed. The absence of the similar fluorescence shows that the quantity of the molecules which are capable of fluorescing is less than 10^{13} to $10^{14}/cm^{-2}$, which at an altitude of 10^7 cm of the homogeneous atmosphere means, that their density near the surface is less than 10^6 to 10^7 cm^{-3} [Urey (Ref. 113)].

10. The extremely rarefied lunar atmosphere should be partially ionized,and therefore its presence can be studied by observations during a period when the Moon is located between the Earth and the cosmic sources of radio emission. Observations of occultation of the Cancer nebulae were conducted in 1956 by Elsemore (Ref. 68) at a wavelength of 3.7 meters. These gave the deviation of the observed duration of occultation from the theoretical at 0.4 ± 0.26 minute. If you consider that the deviation is actual, this indicated the concentration of electrons at the surface of the Moon to be 10^3 cm^{-3}. Since 10^{-3} molecules should be ionized, this corresponds to a density of the molecules 10^6 to 10^7 cm^{-3}, i.e., 10^{12} times less than the density of the near Earth atmosphere.

Nevertheless, the contemporary data concerning the rate of the gas loss by the Moon compels us to doubt that the Moon has even a rarefied atmosphere.

11. It was once considered that the Moon could possess an atmosphere of heavy inert gases - krypton and xenon formed during radioactive

processes and under the effect of cosmic rays. As Urey (Ref. 113) indicated, in the conditions when the layer of dissipation (exosphere) is siutated at a considerable altitude from the surface and is very hot, even these gases are quickly lost. But when the exosphere adjoins the surface, and has as in the last case a comparatively low temperature, the dissipation continues slowly and there is a possibility for the existence of an atmosphere with a density not greater than 10^8 atoms per 1 cm^3.

However, the calculation of the ionizing activities of ultraviolet solar radiation which was conducted by Öpik and Singer (Ref. 99), has shown that the repelling effect of the positively charged lunar /114 surface on the ionized atoms of krypton and xenon accelerates this process to such a degree that the period of semi-decomposition and the rarefied krypton-xenon atmosphere is on the order of 10^3 years. The unpublished data by these authors which was quoted by Watson, Murray and Brown (Ref. 115) give the period of photodissociation and the loss of water vapors in the atmosphere of the Moon as being of a magnitude on the order of one day.

12. The most effective mechanism which removes the lunar atmosphere is evidently the solar wind created by a proton flux, which travels with a velocity of about 10^8 cm/sec. At a density of the flux of 10^2 protons/cm^3, the period of semi-decomposition of the atmosphere is on the order of one day [Harring and Licht (Ref. 77, 78)].

At the same time, the rain of the proton fluxes on to the Moon leads to the existence of a constant atmosphere consisting of neutral hydrogen with a density of about 10^5 atoms/cm^3.

13. The discharge of gases in the Alphonsus crater observed by N. A. Kozyrev (Ref. 20) shows that there still takes place the process of degasification of the lunar interior, i.e., the process of discharge of those volatile substances which went into the composition of the Moon during its formation and which have basically already been discharged for the most part 3-4 billion years ago during the time of the maximum heating of the interior. According to the spectral photometric investigation by A. Kalinyak (Ref. 16), the radiation band photographed by Kozyrev apparently belongs to the molecule C_2 fluorescing under the action of solar radiation just as it occurs in the heads of comets. Although the unique observation of Kozyrev gives no means of judging the frequency and the abundance of such discharges, it is possible to doubt the fact that in the presence of solar wind, they are able to create a lunar atmosphere with a density of even 10^6 to 10^7 particles/cm^3.

14. It was considered until recently that the absence of lunar atmosphere made the existence of ice on its surface impossible. This is undoubtedly true for the parts of the surface illuminated by the Sun where ice can exist only at a certain depth in the form of an "eternal frozen state", protected by a blanket of dust. However, according to Watson, Murray and Brown (Ref. 115), there exist places on the Moon which have never been illuminated by the Sun, where there is a possibility of unprotected ice accumulations. In the lunar polar regions, the mountain slopes turned toward the poles are never illuminated by the Sun, and therefore these are "cold traps" with a temperature of about 120°K. At such a low temperature, the melting of the ice occurs extremely slowly, and on the other hand the accumulation of ice deposits in such traps can occur since a perceptible portion of the water vapors emitted in any spot on the lunar surface should fall into these "cold traps" prior to/115 being "blown off" by the solar wind. Since we observe the Moon from the same direction along which the Sun shines during the full Moon, these cold traps are not accessible for observation from the Earth.

The calculations by Watson, Murray and Brown require a revision, since the illumination of the "traps" by light reflected from the surrounding mountain summits and other illuminated areas of the surface, as well as meteoric erosion, are not taken into consideration.

III. Photometric Data

15. During the full Moon, the lunar disc has an almost uniform brightness (if you disregard the spots caused by the presence of regions with a varying reflecting capability), i.e., it does not display a systematic darkening toward the edge. Under such conditions of illumination, the smooth sphere shows sharp drops in brightness toward the border of the disc. As Galileo noted in his "Dialogues" concerning the two systems of the universe, "the uniform brightness of the full Moon indicates the fact that it has an extremely "rough" surface. Along its border as well as about the center of the disc, the visible brightness is created by a large number of elements from the surface. These elements are equally diversely oriented in relation to the solar rays which are illuminating them.

The uniform brightness of the full Moon to which Galileo turned his attention was only comparatively recently supported by photometric measurements [A. V. Markov (Refs. 25, 26).

16. Approximately during the full Moon, the brightness of the lunar surface is at its maximum [N. P. Barabashev (Ref. 59)]. It decreases rapidly with the increase of the phase angle (i.e., the angles between

the bearings toward the Sun and toward the Earth), and then the decrease slows down [V. V. Sharonov (Refs. 48, 49)]. This indicates the fact that the surface is not simply rough, but deeply "eroded". On it there exists numerous depressions or gaps between the block masses or dust particles (concerning the dimensions of which it is impossible to judge on the basis of one set of photometric data). At such a proximity of the direction of illumination and observation, which occurs during a full Moon, the shadows are practically invisible, but they become visible at the slightest variation of these directions, and this decreases the overall brightness of the surface.

The presence of a sharp illumination maximum during the full Moon is observed at all lunar formations in spite of the diversity of their structure and reflecting characteristics. These are thereby independent of the position of these objects on the lunar disc: at the bottom of the craters, by the seas, on continental areas, and near crater rays. In the case of the latter, the maximum luminosity is particularly sharp [V. A. Fedorets (Ref. 44), Van Diggelen (Ref. 62)]. This testifies to the fact that the microrelief of the lunar surface is similar throughout. The /116 similarity with the photometric characteristics of the lunar surface is displayed by models of surfaces, 65 to 75% of which are covered with holes, the depth of which is greater than their diameter (spherical depressions, as well as spherical protuberances do not give any comparison.) It is interesting to note that the best coincidence with the photometric curve of the lunar surface is given by a curve for lichen (Cladonia Rangiferina). This graphically attests to the extreme roughness of the lunar surface.

17. The great brightness of the Moon, particularly under observations with a telescope, creates a deceptive impression of its whiteness. (In the past this has led to falacious ideas that the Moon is covered with ice or snow. This contradicts the dominating physical conditions of the Moon.) Indeed, the Moon reflects only 7% of the light falling on it. The albedo of the Moon's surface in the optical range amounts to 0.05-0.09 for the dark planes (seas), 0.1 for the continental areas, and 0.17-0.18 for the brightest formations such as the rays of the craters or the bottom of the crater Aristarchus. The limits of the change of albedo on the opposite side of the Moon do not differ appreciably from the corresponding values for the visible part of the surface [A. V. Markov (Ref. 27)]. Thus, the reflective capability of the lunar surface is very low; i.e., the Moon is almost black.

In the radio frequency range, in the 1 cm waveband, the surface of the Moon also has a low albedo, about 0.07 [Hughes (Ref. 79)].

18. The photometric investigations of the lunar surface were encumbered by the circumstance that for any detail the angle of observation changes owing to the libration of the Moon, but only within narrow limits.

The angle of illumination of the detail changes within wide limits with changes in lunar phase. As Minnaert (Ref. 93) already indicated, 20 years aso, the application of the principle of the mutual substitution of angular alignments of the illumination and observation can prove to be a substantial help. Unfortunately, this principle is little utilized for the photometric investigations of the Moon [N. P. Barabashev and V. I. Yezersky (Refs. 3, 4)].

19. In spite of the significant distinctions in brightness of the individual areas of the Moon's surface, color differences on it are practically non-existent [L. N. Radlova (Refs. 33, 34), V. V. Sharonov (Ref. 50)]. The light-color diagrams for the Moon's surface are not similar to analogous diagrams for all known models for the Earth [N. N. Sytinskaya (Ref. 40)] since the limits of the changes of the albedo and the indicators of color for the Earth's formations are significantly wider than those for lunar objects. Apparently this monochromatic condition is caused by the influence of external cosmic factors on the Moon's surface (meteor erosion, the influence of corpuscular fluxes and cosmic rays). /117

Luminescence of the features on the Moon's surface under the action of ultraviolet and corpuscular solar radiation is theoretically possible, but at present, it can not be considered as being conclusively established although certain observation data is available [Dubois (Ref. 64), N. A. Kozyrev (Ref. 19)].

20. The light reflected by the Moon's surface is partially polarized. The measurement of the magnitude of polarization [Lyot (Ref. 89, Dollfus (Ref. 63)] has shown that its maximum occurs at the $\pm 90^{\circ}$ phases of the Moon, i.e., about the first and last quarters. For the dark areas (seas) this maximum amounts to 12-16% and for the brightest areas (the inner parts of the bright craters) about 5%. Furthermore, the plane of polarization is parallel to the equator of intensity ("positive" polarization). During the 0° phase (full Moon), polarization does not occur; and during the phases from 0 to 23° the plane of polarization is perpendicular to the equator of intensity, i.e., the polarization is "negative" with the extreme value of about 1°. It is possible that the "negative" polarization is explained by the repeated reflections from diversely oriented areas [Öhman (Ref. 96)].

The observations in the blue and in the nearby ultraviolet regions of the spectrum have shown greatly increased polarization of lunar light in them [Gehrels (Ref. 71)].

Laboratory measurements of the polarization of light reflected by Earth formations [Lyot (Ref. 89), Dollfus (Ref. 63)] have shown that the best similarity with the curve of polarization for the Moon (and in both

a qualitative and quantitative relationship) is given by curves for dark volcanic ash.

IV. Radio Location Data

21. If, relative to radio waves, the Moon were "white" and so rough regarding light waves, then all parts of its disc would strongly reflect radio waves toward the Earth. The difference between the distances to the center and to the borders of the disc (equal to the radius of the Moon R) would lead to a duration of the reflected signal of $t_0 = 2R/c = 11.6$ milliseconds. However, in the first experiments on the radio location of the Moon, the duration of the echo signal was equal to about 1 microsecond; i.e., it was found that reflection occurs only from the central part of the disc with a diameter of about one third the general diameter. Such a duration of echo signal was observed with wavelengths of 2.5 meters [Evans (Ref. 67], 1.6 meters [Trexler (Ref. 112)] and 10.5 centimeters [Japlee and others (Ref. 80)]. This shows that relative to radio waves, even those of 1 decimeter, the surface of the Moon appears to be quasi-smooth and gives an almost mirror-like reflection. This result agrees with the hypothesis that the surface consists of /118 small particles which resemble sand [Grant and Japlee (Ref. 75)] because in the conditions on Earth during the location of dry sandy deserts from the air, a mirror-like reflection is observed, i.e., a bright spot exactly under the radio locator. This characteristic of the Moon of giving a "mirror" reflection has been utilized for intercontinental radio transmissions.

The analysis of the echo signal showed that after the arrival of the reflection from the central high-light, which has the greatest amplitude (usually several times greater than the noise level from the limb), there fall on reflections from the side high-lights which interfere with one another. These high-lights are formed by the non-horizontal surface areas which are perpendicular to the line of vision and thus give a mirror-like reflection to the Earth. The dimensions of the region from which the side high-lights come show that there is a considerable number of sections on the lunar surface with inclinations up to 5° to the horizontal, but only a small portion of the surface has a high degree of pitch. This portion is so small that the reflections from these areas are often lost among the noises. The application of the present computers of calculation has permitted the separating of signals which are much weaker than the noise. This made it possible to receive the entire echo from the Moon with a duration of 11.6 microseconds [Pettengill (Ref. 30)]. An analysis of the curve of dependence of force of the reflected signal-the angular distance from the center of the disc, showed that along with the "mirror" reflection, from the series of large areas, the lunar surface

also gives a diffuse scattering of radio waves. The curve consists of two parts: at first sharply falling according to the law of exponents (-10.5 sin A), where A is the angle between the incident ray and vertical at the point of incidence or the horizontal distance from the center of the disc; and a more gently sloping part decreasing as $(\cos A)^{3/2}$. The character of the entire curve can be explained on the supposition that 92% of the surface of the Moon is made up of "smooth" area, covered by particles with dimensions less than the length of the radio waves, and 8% is made up of "rough" parts with details larger than the wavelengths. The "smooth" part is mainly seas and the "rough" parts are rocky areas.

22. The time of arrival of the reflected signals, measured with extremely high accuracy (one tenth of a millisecond), gives us the distance from the disc's center of that circle where the reflecting area is located. Simultaneously, the doppler measurement of wavelength which is distinguished, up to a magnitude of $2 \cdot 10^{-10}$, from the wavelength, permits to the observer a very accurate establishment on the surface of the planets a line with even departure velocity or approach velocity. The intersection of this line with the circles gives the location of two symmetrical areas relative to the equator from which the given signal is obtained. The energy of the signal and the degree of its polarization give information concerning the reflective surface. At present, the radio location topographic chart of the Moon with a detailed /119 resolution of up to 100 kilometers [Pettengill (Ref. 100)] has been drawn up according to this principle. It is true that on this chart, the lunar disc appears is if "folded in half" along the equator. The acquisition of similar charts under various librations of the Moon will certainly permit the construction of a full (not folded in half) radio location chart of the lunar disc.

The radio location method has great possibilities since there is almost no limit to its resolution capacity during the increase in power of the locators. The resolving capability of the time frequency analysis is not dependent upon the distance to the object; therefore, its advantages should be disclosed particularly during the study of remote planets.

V. The Temperature of the Surface

23. The radiometric observations have shown that the temperature of the surface of the Moon rises to approximately + 100 to 120°C in the Sun and falls to - 150°C and even lower in the unlighted (night) hemisphere [Pettit and Nicholson (Ref. 102)].

The distribution of temperature along the Moon's disc during full Moon is such that the isotherms appear to be concentric circles. This indicates the fact that the temperature very rapidly reaches an equilibrium value corresponding to the given height of the Sun above the horizon. Hence, the conduction of heat of the surface layer which determines the outward flow of heat deep within, is very small. This is also indicated by the absolute value of the temperature in the Sun. This value is very close to the equilibrium for a surface which is practically non-conductive of heat. The surface gives off all of the received heat in the form of radiation.

24. The measurement of the temperature of the surface area (close to the border of the disc) during the lunar eclipse in 1927 showed a very sharp decline during the penumbral eclipse which corresponds to the specific eclipse of the Sun for the studied area of the Moon's surface. Within the time of passage of the Earth's shadow through this area, its temperature dropped very little, and then with the convergence of penumbral arrival it rapidly rose to the previous value [Pettit and Nicholson (Ref. 102)]. During the eclipse in 1939, with a more precise technique Pettit measured the temperature at the point near the center of the lunar disc. The curve of the change of temperature was found to be similar to the curve in 1927 [Pettit (Ref. 101)].

According to the data concerning the temperature change with lunar phase, at the surface idealized in the form of a planar semi-space they usually get a numerical value of the magnitude $(k\ c\ \rho)^{-1/2}$, where k is the heat transfer, c is the heat capacity, and ρ is the density of the /120
upper area of the lunar surface. Since this value proves to be very large (about 1000), then when $\rho = 2.0$ and $c = 0.2$ (as with rocks on Earth), the heat transfer k proves to be exceptionally low, on the order of $2 \cdot 10^{-6}$ [Wesselink (Ref. 116)]. Such a low heat transfer was not observed in any of the known terrestrial specimens, some of which were exceptionally porous. Only the heat transfer of powders in a vacuum approach this value [Smolukhovskiy (Ref. 111)].

A low value for the dielectric constant of the lunar surface was obtained from radio observations. This implies a low density for the surface layer of 0.4 - 0.5 gm/cm^3, i.e., much lower than that of Earth's formation [A. E. Salomonovich (Ref. 37)]. Since the surface layer itself consists of dust, its density must be exceptionally loose. Such a surface area structure explains its low heat transfer: the small dust particles in the vacuum transfer heat only across small areas of the contact of individual particles.

The thickness of the surface layer that has low heat conductivity is not the same in various regions of the Moon. Measurements of the

thermal flux which is emitted by the Moon during a lunar eclipse show that the temperatures in various spots fall unevenly. Craters with rays (young formations) cool off more slowly than the regions surrounding them. Sinton (Ref. 38), during the eclipse of September 5, 1960, measured the flux of heat on a wavelength 8.0-9.5 μ along the chord crossing the Tycho crater (the center of the largest ray system). During the eclipse, the temperature in the crater remained 30-40° higher than the surrounding regions. This result can be explained by the assumption that the bottom of the Tycho crater is only partially covered by a dust layer and that in certain spots on the surface rocks emerge. (By calculation, 89% is a dust layer and 11% is rock.) Solid rock formations possessing a higher heat transfer during the eclipse continue to radiate the heat coming from the deeper layers. Another explanation amounts to the fact that the thickness of the dust layer within the crater must be about 0.3 mm, and much thicker outside it. The origin of the dust blanket apparently is connected with the bombardment of the lunar surface with meteorites. The fact that this blanket is so thin is in agreement with the relatively young age of the Tycho crater.

Analogous results were obtained by Saari and Shorthill (Ref. 104) for the craters Copernicus, Aristarchus, and Kepler, as well as Tycho.

Evidently, a model of the non-homogenous lunar surface, which would be partially covered by a layer of dust of varying thickness and partially by protruding solid rock formations, is close to actuality. Such a model would agree with the data concerning the surface temperatures much better than the single layer or the solid double layer model.

25. In recent years, measurements of the Moon's specific /121 thermal radiation have been conducted. The depth penetration of electromagnetic waves is two to three times greater than their wavelength; radio emission from different layers beneath the Moon's surface is therefore measured during observations on various wavelengths.

The observations of radio emission give the brightness temperature, i.e., the temperature of the dark body radiating with the same intensity in the area of the spectrum which is under study. In the dark body, the brightness temperature coincides with the true temperature. Thanks to the fact that the Moon, in the radio frequency range as in the visible region, possesses a very small reflective capacity for the center of the lunar disc (in a direction which is normal to the surface), the brightness temperature exceeds the true temperature by only 2 to 3%. However, since the reflectivity increases in the case of an inclined direction, the brightness temperatures drop at the border of the lunar disc. In other words, the form of the brightness temperature isotherms usually in the capacity of radio observation results, differ essentially from the form of the true isotherms.

The measurement ot radio emission on wavelengths from 75 cm to 10 cm originating at depths which exceed the wavelength approximately on the order of one, indicated that the temperature there remains almost constant, during all lunar days and equals 230°K [N. L. Kaydanovskiy, M. T. Turusbekov, S. E. Khaikin (Ref. 15) and Seeger (Ref. 106)]. This temperature stability at a small depth is more independent proof of the very small heat conductivity of the surface layer.

On wavelengths less than 3 cm, originating at depths up to half a meter, a change in temperature relative to the lunar phase has been observed, whereby these changes increase with decrease of wavelengths [M. R. Zelinskaya, V. S. Troitskiy, and L. I. Fedoseyeva (Ref. 14); N. A. Amenitskiy, R. I. Noskova, and A. Ye. Salomonovich (Ref. 1)]. On a wavelength of 1.5 mm, the temperature fluctuations are close to those which are observed in the infrared region [Sinton (Ref. 109)]. On a wavelength of 1.5 mm, as well as in the infrared rays, the temperature maximum of the central parts of the disc appears during a full Moon, and on wavelengths of 12 to 16 mm, the maximum is delayed by approximately 3 days; i.e., it occurs during a phase shift of 35 to 40°.

26. Because of the fact that the axis of rotation of the Moon is almost perpendicular to the plane of the ecliptic, the polar regions of the Moon are always less heated by the Sun. Therefore, at such depth where the temperature remains constant, the latter should decline substantially from the equator toward the poles. Observations with large radio telescopes on ultra short waves, when the dimensions of the antenna beam widths are less than the dimensions of the lunar disc, make it possible to measure the temperature distribution along the lunar disc. The shape of the isotherms, obtained at a wavelength of 8 mm, /122 confirmed the decrease of temperature toward the poles.

Near the equator, the temperature is approximately -40°C at a depth where the temperature is constant. Near the poles, the temperature might drop by 100 to 200°. A similar latitudinal change of the "stationary" temperature of the surface layer should be accompanied by a corresponding change of temperature distribution along the Moon's radius from the equator toward the poles.

VI. The Structure of the Surface Layer (Micro Relief)

27. Due to the absence of an atmosphere on the Moon, all of the meteoric particles down to the smallest dust specks, reach and erode the surface without encountering any obstacles, creating a layer which is known as the dust layer. In the terrestrial conditions a dust substance is friable. This may be explained by the fact that dust particles in

the air are always covered with a film of adsorbed molecules. This film impedes the contact and adherence of the dust particles with each other. In a vacuum, when the adsorbed film is absent, adherence (agglomeration) of the dust particles occurs.

The return of dust particles to the lunar surface, after they have been thrown upward during explosions of meteoric particles which impacted the Moon, leads, as a result of this adherence, to the formation of extremely friable, porous structures. These structures should cover not only horizontal sections of the lunar surface, but also slopes.

Because of the high speed with which meteors strike the lunar surface, there occurs in addition to the granulation and scattering of the particle matter as well as the surface itself, also a vaporization of a certain part of this matter. The condensation of atoms and molecules which penetrate into crevices between the dust particles increases the strength of adherence of the particles, and in addition, it imparts to the surface a black color, which is characteristic of porous, heterogenous substances in which diverse molecules are mixed unsystematically. Such a porous, brittle layer, formed under the effect of meteoric impacts, carries in fact those characteristics which were disclosed by observations: extremely low heat transfer, the corrugation of the micro relief and blackness [N. N. Sytinskaya (Refs. 39, 41, 42) and F. Whipple (Ref. 117)].

It should be added that the micro relief of the lunar surface is defined as a characteristic of structures, which are formed during the ejection and adherence of dust particles, as well as by indentations and small craters, created during meteoric impacts and during an ejection of large fragments. (The latter are capable of creating not only indentations but also mounds.)

28. The speed of the meteoric impacts is so great that the mass of the pulverized substance of the surface itself is tens and hundreds of times greater than the mass of particles which impacted. Therefore, the external porous layer consists primarily, not of meteoric matter, /123 but of the pulverized lunar matter. As a result, in the reflective ability of various areas of the lunar surface, (for example, the "seas" and the "continents") the difference in the optical characteristics of the rock formations which make up these areas is disclosed. This also shows that the substance is usually not scattered at great distances.

After the formation of the layer of perceptible thickness, the further impacts of small particles affect only this layer and only the more infrequent impacts of large bodies break up the matter beneath it. Thereby, the speed of the thickness growth in the porous layer gradually slows down and its contents (of meteoric matter) increases. The increase in thickness in the layer can completely cease in the case of negative

over-all mass balance. This can take place during a sufficient quantity of high speed impacts during which a significant part of the matter which is scattered acquires speeds exceeding the escape velocity (2.4 km/sec) and is lost by the Moon. In this case, quasi-equilibrium is reached during which the rate of loss of matter by the Moon is statistically equal to the rate of destruction of the underlying matter over a period of time [Whipple (Ref. 117)]. At present, the scarcity of information concerning the quantity and speeds of meteoric impacts and concerning the processes of explosion and crater formation does not permit the evaluation of the character of the mass balance or the thickness of the quasi-equilibrial converted layers at various conditions. An indirect property indicating the extremely significant role of the meteoric erosion of the lunar surface is the smoothed out nature of the reflief of ancient lunar craters. Since the discussion concerns the differences noticeable in medium sized telescopes, erosion should have destroyed layers with a thickness of hundreds of meters or even greater [Gold (Ref. 73)]. This is in accordance with the maximum evaluations of space density of matter in meteors. However, the hypothesis of Gold according to which all of this dust was accumulated in the lunar depressions forming layers with a thickness of hundreds of meters is not confirmed (see Section 36), and therefore it is extremely probable that the over-all mass balance is negative.

Since the Moon is bombarded not only by small dust particles but also by large meteorites, and even by asteroids and the nuclei of comets, the thickness of the porous converted layer should be extremely irregular. As was already said in Section 24, the floors of young craters, such as Tycho, Copernicus, Kepler and Aristarchus, is covered with a layer about 0.3 mm thick. Probably ancient caverns exist with an especially great thickness of porous matter.

The accumulation of dust particles in the caverns could have occurred due to lunar quakes breaking down the fragile dust structures and shifting them downward along the slope. It is also possible /124
that there are two mechanisms participating. This was proposed by Gold (see Section 36). Although they are not able to explain the shift of dust particles, at hundreds of kilometers, as is necessary by his hypothesis, it is possible that they are suitable for comparatively small distances.

29. The absence of an atmosphere on the Moon also leads to the bombardment of its surface by solar corpuscular fluxes. The mass of matter brought by these fluxes as comparable with the influx of meteoric matter and may even exceed it (10^{10} protons/cm^2 sec or $2 \cdot 10^{-14}$ gm/cm^2 sec). This does not create an accumulation of matter on the surface of the Moon since hydrogen is rapidly dispersed. On the contrary, this creates a surface dispersion analogous to the cathode dispersion since the energy of the protons amounts to 10^4 to 10^5 electron volts. At such

great energy, the dispersion process continues extremely intensively and can play a role which is no less significant than that of the meteoric erosion. The high speeds being acquired by a significant part of the dispersed matter can create a negative over-all mass balance, and the remaining molecules and grains while settling on the surface should impart to it a dark color (together with the condensing molecules which were vaporized during meteoric impacts).

Cosmic rays and also ultraviolet radiation of the Sun evidently play an insignificant role in the reformation of the surface layer of the Moon. The effect of cosmic rays results mainly in nuclear conversions during which unstable isotopes arise in particular (similar to the process taking place in meteorites). The ultraviolet solar radiation probably gives rise to small changes in the optical characteristics of the surface layer. These changes are analogous to those that appear on Earth when mountain tops are illuminated by the Sun.

VII. The Relief of the Moon's Surface

30. The appearance of the Moon's surface changes remarkably with changing conditions of illumination. At about full Moon, the visibility of details is determined by differences in their brightness, i.e., practically by differences in their coefficients of reflection. However, since close to full Moon, the brightness is strongly dependent upon the phase angle and thereby is somewhat different in regard to different details, the appearance of the Moon's surface changes even with small variations of this angle. Besides, the conditions of illumination and observations of any area of the Moon's surface are also dependent upon the phase of libration in latitude and longitude. The variety of illumination conditions arising from this and the observations is almost impossible to calculate. This has led and continues to lead to fallacious conclusions concerning changes on the Moon's surface.

On surface areas being observed under conditions of indirect /125
illumination, the appearance of details due to variations in their brightness is placed second in comparison with the appearance of surface relief due to shadows cast by the slopes that are turned away from the Sun and partly according to the increased brightness of the opposite slopes. With the approach of the terminator to the area being observed, the more gentle slopes, if they are only sufficiently drawn out in a meridional direction, begin to cast noticeable shadows, changing the appearance of the surface. In the case of very gentle slopes, the penumbrae, i.e., the regions only partially darkened, play a substantial role.

Kuipers descriptions of the lunar surface (Refs. 17, 85, 86) are considered to be the best at the present time. His "Atlas of the Moon" (Ref. 87) introduces a systematic selection of lunar photographs under five different conditions of illumination. These observations were taken with the help of the world's best modern telescopes.

31. On the surface of the Moon, the brighter regions thickly dotted by craters and the dark plains with a fare more level surface stand out. The first are conditionally called "continents" and the second "seas". As was already stated (see Section 5), the surfaces of the seas are on the average 1.5 km lower than the continents.

32. Hypsometric data concerning the relative heights of the separate mountain formations are still extremely fragmentary and inaccurate. Presently the accumulation of this data is intensive in the USA, England, and France.

It is assumed that one of the summits in the Leibnitz mountains at the border of the lunar disc has the greatest altitude, equal to 9 km. The western border of the bank of the Newton crater has an elevation of more than 7,200 meters above the surrounding area.

According to measurements of the lengths of the shadows conducted by Medler in the beginning of the 19th Century encompassing 1,095 summits (because of the absence at that time of accurate charts of the Moon the identification of measured summits was frequently almost impossible), the following distribution according to altitude [J. Frantz (Ref. 45)] was obtained:

6,000-7,000 meters	6 summits
5,000-6,000 meters	21 summits
4,000-5,000 meters	82 summits
3,000-4,000 meters	184 summits
2,000-3,000 meters	289 summits
1,000-2,000 meters	320 summits
less than 1,000 meters	192 summits

Such measurements met difficulties connected with the non-horizontal nature of the surface on which the shadow falls. For the calculation of this, the study of the velocity of the displacement /126 of the shadow is necessary. This is accomplished now with the help

of a slow-motion filming [MacMath and others (Ref. 92); Kopal (Ref. 21, 84, and others].*

33. The most characteristic formations on the Moon are the craters and the circuses-circular depressions surrounded by a gently sloping annular bank. The craters are those depressions with a concave cup-like floor, and the circuses are the depressions with a planar smooth floor usually as dark as the surface of the seas. The circuses are found only among the large formations with a diameter of 10's and 100's of kilometers.** The number of craters and circuses increases with the transition from large to small down to the limit of the resolving power of the telescopes (depressions with a diameter of 0.5 to 1 km). The number of very small craters comprises hundreds of thousands. There is no doubt that the number of smaller craters which are inaccessible to present telescopes is still greater.

The statistical investigation by Young (Ref. 83) has shown that the distribution of all craters according to their dimensions is close to the law of the reciprocal power $N \approx Ax^{-m}$ (x is the diameter of the crater) whereby $m \approx 2.5$.

34. Practically all of the explorers (with the exception of Gold, see Section 36) believe that the seas are evacuated areas which were at some time filled with molten lava. The small number of craters in the seas is explained by the fact that we observe only those which were formed after the hardening of the lava. With favorable conditions of illumination, on certain areas of the seas, the annular crests and the bright circles become visible. All the data indicate the fact that these are traces of craters which existed before the appearance of the seas and which are partially or fully filled with lava. In exactly the same manner, the individual mountains rising above the surfaces of certain seas (for instance Mare Imbriem) evidently are the remainders of a previous relief which remained above the level of submersion.

In the presence of very oblique illumination on the surface of the seas, the gently sloping banks with an altitude of 100 to 200 meters and extending to 10's and 100's of kilometers become visible. The slopes of

---

*The "Astronomical Contributions from the Manchester University" Series III, No. 66, 67, 70, 71, 72, 73, began to publish in 1959 a series of papers entitled "Topographic Explorations of the Moon", which contain results of hypsometric measurements of separate sections of the lunar surface.

**Certain authors, i.e., A. V. Markov, (Ref. 27) refer to all large craters, regardless of the shape of the crater's floor as circuses.

the banks are inclined to the horizontal at an angle of only 1 to 1.5° [van Diggelen (Ref. 61)]. It is considered that the banks arose as a result of the lava's motion and compressions occurring during its cooling.

Certain seas have a circular form and are surrounded by mountain chains which as a rule become detached from the side of the sea with a precipice having an altitude of 2 to 3 km. (Individual summits rise to above 4 km.) These mountains are not of the folded type but are /127
deposited formations. The largest circular sea - Mare Imbrium - reaches 1000 km in diameter, and the small circular seas with a diameter of 300 km and less form a continuous transition to the large circuses [Vegener (Ref. 6)]. The seas with an irregular form of the type Mare Nubium are not surrounded by mountain chains and evidently are covered with lava which over flow from the circular seas [Kuiper (Refs. 85, 86)]. In the irregular seas, judging from the abundance of half-submerged craters and other mountain formations, the layer of lava is thinner than in the circular seas where submerged formations are more infrequently observed and thereby only in the outlying sections.

35. The lunar circuses, i.e., the annular plains with dark very planar floors (in the case of very large circuses, for example Ptolemaeus, these are protuberant in relation to the curvature of the lunar surface) are craters filled with lava. These circuses are situated usually in seas or in the vicinity of seas and are filled with lava from them, i.e., judging by the period of their formation, these are undoubtedly of a "marine" type. However, there also exist circuses situated far from the seas or even at elevated places (the latter is only conjectural because of the inaccuracy and incompleteness of the hypsometric data). In these cases it is necessary to assume that the lava arose from beneath. The crater Vargentine is the only one of its type represented as a table mountain since it is filled with lava considerably higher than the surrounding area to a height of the lowest part of the bank [Kuiper (Ref. 86, Dollfus (Ref. 13)].

It is possible to trace the continuous transition from circuses with dark flat floors toward the half-submerged craters at the bottom of seas, from which only low circular crests or parts of crests remain, rising above the surface of the seas, toward completely submerged craters which are scarcely distinguished in the form of bright circuses only. According to the period of their formation, the submerged and semi-submerged craters are of a "pre-marine" type.

At the bottom of certain circuses, for example Ptolemaeus, along with the miniature craters of a normal shape which were formed after the hardening of the lava there are very gently sloping depressions, noticeable only under the most oblique illumination. These depressions are undoubtedly

the traces of overflowing craters which were formed during the last stage of hardening.

36. According to the hypothesis by Gold (Refs. 9, 73, 74), the dark floors of the seas and of the circuses is covered not with hardened lava as is assumed by a majority of specialists, but by a thick layer of dust, perhaps several kilometers thick. From the viewpoint of future landings on the Moon, the assumption that such a layer may be similar to quicksand [Gold (Ref. 74)] is of great concern, and has been /128
repeatedly discussed in the literature. The main difficulty for Gold's hypothesis is the problem of the shifting of dust at large distances along the lunar surface in the absence of air and water flows. Actually, according to these assumptions, the particles of dust which were formed on elevated parts of the Moon's surface-"the continents", should be shifted on the order of hundreds and thousands of kilometers in order to reach the sea depressions and remain there. Gold considers that the electrostatic charges formed during the photoionization of the ground must impart motion to the dust particles. He considered two mechanisms of motion-jumps and sliding. In the first mechanism, the repulsion of similar charges is examined which, according to the calculations of Gold is capable of breaking away individual dust particles from the surface layer and forcing them to jump. The change of position of dust particles downward along the slope should be the statistical result of the number of such jumps. In the second mechanism, the suspension of the charged dust particle at a short distance (about 1 mm) from the oppositely charged surface is assumed and in the case of a non-horizontal nature of this surface, sliding downward along the slope will occur.

However, the surface of the continental regions of the Moon is too uneven for these mechanisms to give the effect which Gold has assumed. Neither the jumps nor the sliding can take the dust particles out of a crater depression or from regions surrounded completely by crater banks. Nowhere can Gold's assumed "dust rivers" "flowing" slowly toward the seas be seen. Therefore, his hypothesis, although it received wide acclaim, is not shared by specialists.

37. At the present time, the nearly century-long argument among the supporters of the "volcanic" (more accurately-endogenic) and the supporters of "meteoritic" or "impact" origin of lunar craters still continues. The majority of explorers of the Moon now recognize the formation of the craters as a result of the bombardment of the lunar surface by bodies with diameters of sometimes tens of kilometers. This occurred on the background of radio-active heating and melting of the lunar interior, and because of this they were accompanied by lava flows. The first stages of this bombardment are related to the formation of the Moon itself (see Sections 49, 50, and 51), and the last stage, which is still in progress is related to collisions of asteroids and comet nuclei

with the Moon. These asteroids and comet nuclei fly into the internal part of our planetary system. Öpik (Ref. 98) compared the surface density of the craters at the Mare Imbrium with the present frequency with which large bodies-small asteroids and nuclei of comets, impact our Earth. Öpik (Ref. 97) evaluated this earlier. He came to the conclusion that these data are in mutual accord in the supposition that the formation of the craters in the Mare Imbrium occurred during /129
the last billion years almost periodically. This is confirmed by the geological investigations of the crater-like structures on the territory of the USA which indicate that the rate of the Earth's bombardment within the last half-billion years, was probably the same as the rate of bombardment of the lunar surface which has created post-marine craters [Shoemaker and others (Ref. 108)].

The distribution of impacting bodies in respect to their masses can be derived from the distribution of lunar craters in respect to their dimensions, which was obtained by Young (Ref. 83). The thus derived distribution of crater forming bodies in respect to their masses is in agreement with such a distribution for asteroids [Jaschek (Ref. 81)].

A number of small craters of a "secondary" origin exist side-by-side with the craters formed through impacts of extra lunar bodies, at the banks and in the vicinity of large craters. These small craters were formed during the impact of returning fragments of the lunar surface which were thrown upward during the formation of these large craters.

The energy of the impacts which formed the smallest lunar craters with diameters of about 1 km, was equivalent to the explosion of 1 to 2 megatons of TNT; i.e., it amounted to about 10^{23} erg. (This was determined, for example by a comparison of the meteorite crater on the Earth in Arizona with the craters formed during American underground explosions of atomic bombs [Shoemaker, (Ref. 107)]. The energy of the explosions which formed the largest lunar craters and the circular seas should have reached 10^{30} to 10^{32} erg.

About half of this energy was converted into an energy of seismic tremors the intensity of which during the formation of large craters exceeded the energy of the cataclysmic earthquakes on Earth. Because of the small dimensions of the Moon, the surface waves, the attenuation of which is very small, could have been focused in a diametrically opposite point creating what seemed like a second epicenter (Kopal, 1959). The "meteorite" hypothesis of the formation of lunar craters in its present form explains all observed forms of the lunar relief [B. Yu. Levin (Ref. 23)].

The "volcanic" hypothesis is primarily presently supported by several geologists (Sperr, Khabakov, Byulov) who are not concerned with the physical foundation of the processes which they proposed and are not taking into consideration the development of the lunar interior. This leads them to unrealistic concepts.

38. The craters throughout the surface of the Moon are not distributed randomly as it should be since in the seas, only a small number of craters, which have been formed after the hardening of the lave, are visible. Furthermore, the deviation from the random distribution of the observed crater centers is stipulated by the fact that the formation of a large crater nullifies the smaller craters which earlier existed in this place. Finally, there are small craters of a /130 "secondary" origin in the vicinities of the large craters (see Section 37).

If we exclude the secondary craters, or take into consideration the influence of the above indicated factors, then within the boundaries of the homogeneous regions of the lunar surface-continents and seas-the distribution of craters is at random. On the surface of Mare Serenitatis where the impositions are nearly absent, the distribution satisfies the Poisson law [Arthur (Ref. 55)]. For Mare Imbrium, small deviations are noticed that are evidently related to the presence of secondary craters which were formed by ejections from the Copernicus crater. Due to the absence of statistical methods for an analysis of the crater distribution in regions where the impositions are substantial, Arthur arranged purely random distributions of superimposed circles with various diameters and showed that these models do not differ from the crater fields in the Moon's continental regions.

An investigation of the surface density of large craters (with a diameter greater than 20 km) at individual areas of the continents, which were conducted by Green (Ref. 76) led also to the conclusion that they are distributed at random.

39. The cases of imposition of one crater on top of another, which are especially frequent in the continental regions enable us to determine their relative age: the superimposed craters are younger than the craters beneath them. The comparison of their outer form shows that the young craters have a more abrupt relief at the floor and banks while the old craters relief is rather smooth. The latter is connected with seismic destructions, by meteoric erosion and erosion caused by corpuscular fluxes. The most abrupt relief can be found in craters with rays which are the most recent whereby, as now disclosed, they are not only relatively younger, but absolutely younger.

The study of distribution and the outer form of the craters which was conducted by many investigators has made it possible to segregate them by their respective ages. The "pre-marine", "marine", and "post-marine" craters, that is, the craters which were formed before the seas, during the seas, and finally, after the seas, are now distinguished [A. V. Khabakov (Refs. 46, 47) with no basis has introduced an intermediate era when the process of crater formations has diminished for some reason.]

40. The lunar craters are extremely planar formations, the depths of which are many times less than their diameter. The apparently great depth of the craters is an illusion related with the fact that we are accustomed to the shadows which are cast by objects when the Sun is high. At the same time a planar crater, illuminated obliquely does not differ at first sight from a deep crater which is sharply illuminated. The images of very steep crater banks and of lunar mountains on many /131 illustrations made even by specialists-astronomers are based only on this illusive impression.

Measurements show that in craters with a diameter of 20 to 30 km the entire depth (counted from the summit of the bank) amounts to a 10% diameter on an average, and in craters with a diameter of 100 km about 5%, and in still larger craters 2.5 to 3% [Ebert (Ref. 66)]. At the same time in the smallest measured craters with a diameter of 2 to 3 km the depth reaches 20% of their diameter [Baldwin (Ref. 57)].

The elevation of the bank rose with the increase of the crater's diameter comprising approximately a permanent portion of its depth (about 1/3 of the entire depth or about 1/2 of the true depth, which is counted from the level of the surrounding plain).

The slopes of the bank are very gentle, whereby the curvature of the interslope decreases with the increase of the crater's diameter. This is evident in the following data of Faut (Ref. 43):

Diameter of the Crater in km	Mean Diameter in km	Number of Craters	Mean Angle of Incline of the Inner Slope in degrees
30	12	113	33.5
30-50	38	14	22.7
50-100	76	27	14.8
> 100	144	8	11.6

The outer slopes are even more gentle and have an incline of 1 to 6°.

The photometric measurements of the profile of the lunar rim during eclipse of the Sun are an independent confirmation of the small curvature in the relief forms of the lunar surface [Fujinami, Ina, and Kawai (Ref. 70)] who showed that a predominant part of the observed slopes is within 3 to 14° for a scale greater than 30 km.

As a comparison, it is necessary to indicate that the natural angle of repose for a fine volcanic ash amounts to 30 to 35° in the case of a deeper detritus. The fastest flowing basaltic lava of the Hawaiian and Icelandic volcanos while cooling forms slopes, the angle of incline of which rarely exceeds 8°.

41. Many craters have a gently sloping central mound sometimes consisting of several elevations.

Among the "youngest" craters (that is, the craters with an abrupt uneven relief) with a diameter of more than 16 km, about 80% have central mounds. In the small dimension craters, the presence of a central mound is, as a rule, uncertain. This evidently is related to the large curvature of their inner slopes. This leads to the fact that at a low position of the Sun, which is necessary for the disclosure of the central mound, /132 the floor of the crater proves to be in the shade cast by the bank. In the case of "young" craters, the height of the central mound amounts to about 1/2 of the true depth of the crater on an average. Its summit raises only in three cases above the level of the surrounding plain but never reaches the level of the bank. In the case of "old" craters with a smooth relief, the central elevations are also smooth, and therefore are only rarely observed.

As a rule, circuses have no central mounds since they are submerged in lava.

There exist miniature craters on the central mounds of certain craters. Although their external appearance indicates a random superposition of the small crater on the central protuberance, the supporters of the volcanic hypothesis on the origin of lunar craters have considered them as important proof in their favor. Baldwin (Ref. 57) after having evaluated the relative area of the central mound found that in a purely random crater distribution, there should be observed about 15 cases of superposition of the small craters on the central mounds and actually there were twelve such cases known at that time. Baldwin's results caused special searches on the Moon for mountains with craters near their summits. This led to the increase of their numbers to 50 [Moore (Refs. 94, 95)]. However, the expanded list included not only the central mounds, but also the lava domes which are of a different nature (see Section 47).

42. The so-called "law of Schreter" is exceptionally important, according to which, the circumference of the bank on the average is equal to the circumference of the depression. The reality of the "law of Schreter" is negated by certain investigators on the basis that it is accomplished only on the average. However, the deviations from it often have a natural explanation. Thus, for example, cases when the circumference of the depression are found at submerged craters (i.e., those having a planar dark floor). Baldwin (Ref. 57) has shown, that based on the Schreter law and considering the width as being equal to 1/4 of the crater's diameter, as well as the average dimensions of the central mound, it is possible to calculate the theoretical dependence of the bank elevation by the observed dependence of the crater's depth on its diameter (see Section 43). This is in excellent agreement with the observed dependence.

In this way, the lunar craters are on the average "0" shaped relief unlike the Earth volcanos which are "positive" relief forms and the Earth calderas which, relative to their nearest surroundings, are "negative" relief forms. (Calderas are dip formations situated on the summit of volcanos and volcanic domes. This is a unique type of volcanic formation somewhat similar to the lunar craters).

43. In the lunar craters relationships exists between the diameter and the depth and also between the diameter and the height of the bank. They are a rolling continuation of analogous relations established for shallow holes and craters which are formed on the Earth from explosions of munitions, bombs and more powerful charges of explosive material [Baldwin (Ref. 57). The interval between the maximum explosion craters on the Earth and the minimum lunar craters accessible to study are covered by the Earth's meteoric craters. (Recently Shoemaker (Ref. 107) established the complete similarity between the Arizona meteoric crater and two craters formed during the American underground atomic explosions.) These relations bear a merely statistical character, and for the lunar and Earth craters the deviations from strict regularity are very great. For the Earth caldera, this dispersion is immeasurably large. /133

The relations between the diameter, the depth, and height of the bank and also the direct study of the outer form of the craters show that all craters form a continuous succession. Furthermore, the circular seas surrounded by mountain ranges form a further continuation of this sequence [Vegener (Ref. 6)].

44. The youngest craters are surrounded by a crown of long bright rays extending (for example, in the Tycho crater) more than a thousand kilometers in individual cases. The rays are visible at about full Moon and disappear during oblique illumination. They are extended across seas

and continents and are undoubtedly deposited formations. Even during extremely oblique illumination, the rays do not cast shadows, i.e., the height of the substance deposit is very small. Photometrical investigations have shown that the rays do not change the dispersing properties of the underlying surface. This also indicates the thinness and scarcity of the substance deposit [N. P. Barabashov and V. I. Yezerskiy (Ref. 4)]. Certain rays are not radially directed but are tangent to the crater or deviate even more from the radial direction. This seemed to create difficulties in the explanation of their ejections from the crater. Giamboni (Ref. 72) on the basis of calculations of ballistic trajectories on the surface of a rotating planet showed that deviations from the radial directions can be explained by the formation of rays at the time of a rapid lunar axial rotation (with a period of 0.5 to 3.5 days). After it was disclosed that the craters with rays are not only relatively, but absolutely young, this explanation was no longer valid. Even if their age does not amount to 10^5 to 10^6 years as Levin (Ref. 23) admits but is 10^7 to 10^8 years as Sinton (Ref. 110) holds, all the same the Moon during their formation did not have a rapid axial rotation.

However, the strictly radial direction of the rays could have arisen only during very special conditions of the ejection of substance: during the ejection of substance in one definite direction with different velocities or with a planar vertical fan of ejections. The inclined shower of ejections or the changing of direction of the "stream" being thrown out should have led to the non-radial direction of the ray, which is the geometric position of the intersections of trajectories of individual particles with the Moon's surface. /134

Certain rays of the Copernicus crater consist of separate successive lines each of which is strictly radial, but their totality is not radial. Apparently each line arose as a result of pulverization during an impact on the surface and scattering along the direction of motion of the ejected fragments of bright substance.

Not all the ejected fragments are frangible; there are those which remained in one piece. One can even see furrows drawn by them [Kuiper (Ref. 86)]. Along the rays, a number of elevated small craters is also observed. This indicates their formation as a result of the falling of fragments ejected during the formation of a large crater. It is possible that the fragments which dug the furrows flew along flat trajectories, and fragments which formed small craters flew along curved trajectories.

Certain young craters, in particular those which were formed on the even surface of the seas (for example Copernicus), are surrounded by a corona of craters caused by radial ejections the extent of which approximates the diameter of the crater. As along the rays with the boundaries

of this corona an increased number of miniature craters is observed. This is due to their formation during impact of fragments which returned to the surface after ejection.

In a number of lunar regions, there exist chains of small craters which are sometimes partially superimposed upon each other. Their formation is evidently related to the ejection of a series of fragments, analogous to those which produced the echelon type rays of Copernicus.

45. On the surface of the Moon, numerous fissures are observed extending in a straight line or with small deviations to distances of 50 to 100 km and more. Their width generally amounts to several km. The majority of the fissures are situated on the surface of the seas and at the floors of craters submerged in lava. The formation of these fissures is evidently related to the cooling and shrinkage of the lava layer. Such for instance is the system of fissures on the floor of the Alphonsus crater which is approximately concentric to its embankment, or the system of echelon fissures along the surface of Mare Tranquillitatis. The latter (like many others) have a planar floor. Kuiper explained this phenomenon with the theory that the craters were filled with liquid lava during the period of fissure formation.

Systems of fissures also exist that are concentric and radial to the circular seas. The formation of these systems of crevices is un- /135 doubtedly related to the formation of the seas. The three arc-shaped fissures stretching along the western edge of Mare Humorum and the radial crevices which emanate from Mare Imbrium incompassing the Alpine valley and the so-called Straight Wall, are examples of this configuration.

The fissures of the above indicated type do not exhibit perceptible differences in the height of both rims. The only exception is the Straight Wall. It consists of an embankment 220 to 370 meters high extending along a distance of 500 km [Ashbrook (Ref. 56)]. However, the wall of the embankment is not vertical as is sometimes claimed in the literature, but consists of a gentle slope, the inclination of which is only 18°.

Finally, a third variety of fissures must be considered. This is exemplified by the system of crevices which surrounds the plateau of Archimedes, which is a large polygonic section near the Archimedes crater, that is, slightly raised above the surrounding surface. Horizontal displacements along fissures are, as a rule, not observed. Only two or three cases of such shifts are known.

46. Besides the narrow fissures on the surface of the Moon, there exist wide furrows and flat valleys. In the central region of the lunar disc, the majority of these formations are oriented radially to Mare Imbrium. The furrows, particularly when they extend along the highest

points of the lunar relief, obviously originated during the formation of the Mare Imbrium when enormous fragments were ejected in a nearly horizontal direction. The radial valleys could have originated either as voids between neighboring radial ridges, or as a result of the accumulation of falling debris deposited at sufficiently low velocity not to produce an explosive action but to create a canyon in the form of a drawn-out crater. It is necessary to note that descriptions of similar formations presented by various authors and the terminology used are quite dependent upon assumptions concerning the mechanism of origination for these furrows and valleys.

47. On the surface of the shallow (submerged) seas and circuses, one notices lava domes in the form of circles or sometimes ovals. At the pinnacles of 10 to 15 of these circular domes, whose diameters are 5 to 10 km, tiny craters are noticeable at the limit of visibility, that is, with a diameter of about 0.5 km. The angles of the slopes of these domes are 5 to 8°, that is, they are similar to Earth volcanos formed by liquid basalt lava. There exists a basis for the assumption that these formations are in fact extinct lunar volcanos of the type found in Hawaii.

The fact that no craters were observed on the pinnacles of some of the domes can be explained on the basis of their possible small dimensions. The statement made by Arthur (Ref. 2) that certain domes are topped by protuberances rather than by craters lacks confirmation.

VIII. Origin, Internal Structure, and Thermal History of the Moon

48. Until the 1940's, during the era of the cosmogonic hypotheses, it was postulated that the Earth and the planets were formed from molten coagula of matter. An identical origin was ascribed to the satellites of the planets including the Moon. The formation of the Moon was explained on the basis of Darwin's hypothesis, although this hypothesis was found to be erroneous long ago. According to this hypothesis, the Moon was separated from the rapidly rotating molten Earth. The thinness of the Earth's core at the bottom of the Pacific Ocean and the absence of a granite layer provided a basis for the assumption that the Moon emanated from a spot in the Pacific Ocean. Darwin's hypothesis was primarily based on his investigation of the equilibrium shapes of a rotating liquid mass. According to his findings, a mass with a sufficiently high rate of rotation could separate somehow into two unequal masses. One of the masses would begin to rotate about the other at a small distance apart. Second, the Darwin hypothesis is also based on his investigations of the resonance fluctuations of a rapidly rotating liquid Earth out of which arose the possibility of a coincidence between the period of its own fluctuation and the period of

solar tides. Darwin's analyses of equilibrium shapes were questioned in the work of A. M. Lyapunov, and at one time the answers were disputable. However, the validity of A. M. Lyapunov's work was established long ago, and Darwin's deductions, as well as the supporting results of Jeans (Cartan) (Ref. 60) and Littleton (Ref. 88) were shown to be incorrect. Darwin's research on the resonance fluctuations was based on the use of linear approximation and did not take into consideration the dissipative forces that occur at large fluctuation amplitudes. As Jeffrey's (Ref. 82) showed, the amplitude of the fluctuations of the liquid Earth could not have exceeded 20 km, i.e., even if they did exist, they could not have played any kind of role in the separation of the Moon. Furthermore, investigations of the ocean floors showed that the Pacific Ocean has the same floor structure as the other oceans. Nevertheless, the hypotheses concerning the separation of the Moon from the Earth to this time continues to be considered by certain authors as though it had a scientific foundation.

49. At present, an overwhelming majority of the investigators in the area of planetary cosmogony consider that all the planets or at least the planets in the Earth's group were formed through the accumulation of cold bodies and particles and were initially cold themselves (O. Yu. Schmidt, Urey, Kuiper, Hoyle, Gold). It is assumed that the Moon was formed in an analogous manner. However, no matter how strange it may seem, the assumptions on the origin of the Moon are closely related to the problem concerning the nature of the dense nucleus of the Earth.

If it is assumed that the nucleus of the Earth consists of iron, then it appears that the Moon, with its small mean density, 3.33 gm/cm^3 cannot contain as much iron as the Earth; i.e., it has a different /137 chemical composition. In this case, it could not have been formed simultaneously with the Earth from the same substance. A supporter of this point of view, Urey (Ref. 52) proposed that the Moon was formed at some distance from the Earth. This provides an explanation for the Moon's different composition. The Moon was subsequently captured by the Earth in the course of a near approach. However, the probability of capture as Urey himself recognizes it, is very small. In addition, it is very unlikely that the Moon traveled in a plane near that of the ecliptic in a manner to be captured into an almost circular orbit.

If, in accordance with Ramsey's hypotheses, the nucleus of the Earth is assumed to consist of silicates which have passed into a metallic state under the influence of high pressure, then the Moon and the Earth should be of similar composition, which would readily explain their simultaneous formation. This leads to the concept that the Moon was formed from a cluster of bodies of various dimensions which surrounded the Earth during the period of its formation.

O. Yu. Schmidt (Ref. 51) assumed that this cluster was formed as a result of inelastic collisions of bodies which took place in the vicinity of the growing Earth and which helped form the Earth. V. V. Radziezsky (Ref. 32) proposed a mechanism for the merging of particles based on the increase of the Earth's mass. As E. L. Ruskol (Refs. 35, 36) showed, the mechanism of O. Yu. Schmidt is fully capable of providing for the formation of a cluster of particles of sufficient mass orbiting the Earth, at a time when the V. V. Radzievskiy mechanism exerts very little influence.

50. The density of the cluster of bodies which was captured as a result of inelastic collisions should have decreased sharply as a function of distance from the Earth; therefore the Moon should have been formed in the dense part of the cluster 5 to 10 times closer to the Earth than its present position [Ye. L. Ruskol (Ref. 36)]. The Moon subsequently shifted to its present position in consequence of tidal interaction with the Earth. The increase in the diameter of the Moon's orbit occurred relatively rapidly at first, but is currently changing at a very slow rate.

This mechanism is in agreement with the commonly accepted theory of the tidal evolution of the Earth-Moon system; but contrary to the Darwin hypotheses, at the beginning of the tidal evolution the Moon was located not within the Roche limit, but rather outside it.

51. The formation of the Moon from a near Earth cluster of particles permits us to conditionally subdivide the history of the surface bombardment of the Moon into the following three immediately consecutive stages:

1) The bombardment in the process of the main accumulation of the Moon which should have ceased within $3\text{-}5 \cdot 10^8$ years after the beginning of the accumulation because of the exhaustion of the material in the "zone of nurishment" of the Moon;

2) The bombardment due to the increase in the lunar orbit radius, which should have ceased within $1\text{-}2 \cdot 10^9$ years due to the Moon's departure from the dense part of the near Earth cluster. /138

3) The still continuing bombardment by bodies that travel along highly eliptical orbits about the Sun (small asteroids and comet nuclei).

During the first stage of bombardment and in the beginning of the second stage, the Moon's interior was still solid. However, the second stage lasted for such an extended period of time that the interior of the Moon had time to heat up and become partially molten (see Section 52). The traces of the first stage are barely preserved. In the beginning of

the second stage, ancient "pre-marine" craters should have been formed. At the end of the second stage, the lunar seas, circuses submerged in lava, and "marine" craters in general should have been formed. In the third stage, the "post-marine" craters, particularly all the craters on the surface of the seas, were formed. At the end of this third stage, - cosmogonically quite recent - the craters with rays were formed.

The differences in the velocities and the nature of the bodies which bombarded the Moon should have created the differences in the character of the craters engendered by their impacts. During the first two stages of bombardment of the Moon, mainly stoney bodies from the near-Earth cluster, possessing low selenocentric velocities reached the surface. Even after acceleration caused by lunar attraction, these bodies impacted the surface with velocities of 3 to 5 km/sec. Since the transition from a partial to an almost complete explosion takes place in this narrow energy interval, the influence of the impact velocity on the structure of the crater is significant. During the third stage of bombardment, the velocities reached meteoric values (up to 72-73 km/sec), with an average value of 20 to 30 km/sec. Thus, the "post-marine" craters would differ somewhat from the "marine" and "pre-marine" types. In addition, differences in the structure of the "post-marine" craters produced by rocky asteroids and glacial nuclei of comets are possible.

The secondary craters were caused by ejected fragments impacting at velocities of 1 to km/sec. The explosive phenomena were affected by the simple compression and displacement of the surface material. Therefore, many secondary craters exhibit a somewhat drawn out shape, rather than being circular.

52. The temperature of the Moon could not have been substantially increased by impacts of bodies which contributed to its formation, because heat due to impact was emitted on the surface and was easily radiated into space. The "initial" temperature of the Moon, which was established at the moment of the practical completion of its formulation, has been determined on the basis of its radiogenic heat. The central regions, which were the first to be formed had sufficient time to reach higher temperatures than the external parts, which had only recently been formed.

The average content of radioactive elements in the Moon as /139 well as the average content in the Earth is assumed to be the same as the average content in meteors. The latter, unlike the Earth's rock formations are specimens of a non-differentiated silicate substance. If the formation of the Moon took place during a period of 3 to $5 \cdot 10^8$ years, then, with the present (low) evaluations of the content of radioactive elements in meteorites, it is found that the "initial" temperature in the center of the Moon amounted to about 500°C.

Calculations of the further evolution of heat of the Moon by radiogenic means [B. Yu. Levin and S. V. Mayeva (Ref. 24); B. Yu. Levin (Ref. 23)] show that within 1 to 1.5 $\cdot$ 10^9 years after its formation, its interior should have been partially melted. Since the pressure in the interior of the Moon is small, the melting signified the transition of matter into a liquid state and should have been accompanied by a differentiation in density and by its chemical composition.

In the case of a greater content of radioactive elements, the melting would have occurred earlier within 0.5 to 1 $\cdot$ 10^9 years after the beginning of the accumulation and would have been, not partial, but complete. Since the evolution of heat and the melting were accompanied by a decrease in density of the matter, this might have led to the instability of the external rigid layer. This layer should have been completely broken and submerged under the floor of lava. However, the dark lava maria occupy only a part of the lunar surface and the bright continental regions probably are remainders of the "initial" matter deposits [Kuiper (Ref. 86)].

In the case of a partial melting, which is more probable, the instability of the external layer should have been significantly less. Therefore, the formation of the lava eruptions ("maria")was not spontaneous but rather was evoked by the falling of large "planetesimals" on the Moon.

The decrease of the over-all quantity of radioactive elements due to their disintegration, their rising to the surface in the course of the differentiation of the interior which facilitated the escape of heat into space, and also a significant loss of temperature during lava eruptions (3 to 3.5) $\cdot$ 10^9 years ago has led to a transition from the Moon's heating to its cooling off. The present distribution of the temperature of the Moon is dependent primarily upon the present distribution of its radioactive elements. Most probably, the Moon is now entirely rigid. Even in the case of a significant content of radioactive elements in the Moon's interior, the external layers down to a depth of 500 to 700 km prove to be rigid.

53. The evolution of heat of the lunar interior should have been accompanied by their degasification. The extent of this phenomenon was dependent on the poorly known content of gases and volatile matter in the primary lunar substance. Because of the small force of gravity on /140 the Moon, the gases are relatively quickly dispersed into space, without forming a stable atmosphere (see Section 7).

The ejection of the gases which was observed by N. A. Kozyrev (Ref. 19) shows that even during the period of heat evolution of the Moon, not all of the gases reached the surface, but part of them still remains in the Moon's interior.

Certain authors, for example Green (Refs. 10, 76), connect the formation of the craters with the degasification of the Moon, considering the former as formations of a type of caldera . However, the form of the craters and the gigantic dimensions that many of them have cannot be explained by such a process. But the existence of small calderae on the summits of lava domes is not excluded (see Section 47).

54. The heat evolution of the lunar interior should have evoked an increase in the Moon's radius of 5 to 15 km (relative to the value of the coefficient of thermal expansion [Mac Donald (Refs. 90, 91)]. The stretching of the external layer by the expanding interior probably was not accompanied by the formation of large fissures of a planetary scale at that time, this external layer still consisted of a poorly consolidated matter deposit. If any fissures were formed regardless, then on the continents there was time for them to fill up during the formation of the craters and of the meteoric erosion, and on the seas they were covered with lava eruptions.

During the cooling stage, the shrinkage of the radius was insignificant since the interior of the Moon today is hot although it is not rigid. Primarily, cooling and shrinkage of the external layers which rested on the slower cooling inner layers, occurred. This should have led and continues to lead to the accumulation of thermo elastic stresses (mainly expanding) which ceased by the formation of internal eruptions. Their formation should be accompanied by seismic jolts.

The absence of a substantial shrinkage of the Moon's radius during the last 2 to 3 billion years is confirmed in view of its surface - because of the absence on it of folded formations and by the almost complete a absence of noticeable horizontal shifts of individual portions of layers [Mac Donald (Refs. 90, 91)]. (Mac Donald believes that the respective shifts go quite unnoticed. However, A. V. Khabakov (Ref. 46) presents as an example the Capella crater divided by a crack along which a shift of 6 to 7 km occurred).

On the present surface of the Moon, the fissures are observed only in the seas and in the lava-filled circuses, i.e., where a stable layer of cooling and hardening lava existed.

The presence of craters, situated in the center of a star-like pattern of fissures or in the cracked areas, and the greatest width of the crevice indicates the formation of these fissures as a result of the impact of a crater forming body into an expanded surface area. Along with this, since the explosion which created the craters damages the /141 entire surrounding ground, cases are possible where the fissures passing through the craters were formed after the crater.

55. Observations with the help of cosmic rockets have shown that the Moon does not have a noticeable over-all magnetic field [Sh. Sh. Dolginov (Ref. 12)]. This agrees with present assumptions concerning the absence of a liquid nucleus in the Moon.

If the Moon contains, as do the meteorites, about 10% iron and if this iron could have flowed to the center during the differentiation and could have formed a liquid nucleus, then with a more probable content of radioactive elements it could have hardened later. If it had remained in a liquid state, it is scarcely probable that convective motions necessary for the creation of a magnetic field, could exist in this nucleus, which is almost void of radioactive heat sources and which is surrounded by a rock cover containing such sources.

However, the low density of the Moon and certain geophysical data are in favor of the hypotheses that in the Moon and Earth, unlike meteorites, practically all of the iron is found in the oxide state [Urey (Ref. 52)]. In this case there is absolutely no basis in expecting the presence of an over-all magnetic field on the Moon.

56. The oblate nature of the Moon's dynamic contour, which is its main deviation from equilibrium can be explained by the fact that the Moon hardened during a state of free rotation [B. Yu. Levin (Ref. 22)]. In such a case, this rotation was decelerated later on due to the dissipation of energy during tidal deformations of the Moon's already rigid body. Hence, the rheologic characteristics of the Moon's substance are such that these deformations do not occur in perfect elasticity (analogous characteristics of the Earth's substance were explained by N. N. Pariyskiy (Ref. 29) during his study of solid influxes in the Earth).

57. Since the pressure in the interior of the Moon is small (even in the center, it amounts to about 50,000 atmospheres, i.e., it corresponds to the pressure in the Earth at a depth of about 150 km), the melting temperature only slightly increases from the surface to the center. Therefore, the partial melting and differentiation should have encompassed practically the entire Moon (with the exception evidently of the "continental" regions of the external layer). According to the geochemical data, from the silicate substance of a meteorite type, it is quite possible to melt up to 15% of the light silicic substance which is similar to the substance in the Earth's core [A. P. Vinogradov (Refs. 7, 8). Therefore on the Moon, the silicic substance with a density of about 2.8 gm/cm^3 can amount to 10 to 15% of its mass. If all of it were raised to the surface, having substituted the initial substance deposit, then the Moon's core would have been formed with the thickness of 80 to 100 km. However

the presence on the Moon of the "seas" and the "continents" forces us /142
to assume that the process of differentiation did not reach completion up to the time of the Moon's hardening.

If the Moon's interior substance was entirely homogeneous and located under isothermic conditions, then its density would have increased to the center by 2 to 5% relative to the pressure increase. However, the temperature increase with depth is greater than the compensation of the pressure increase so that a decrease in density with depth would have been found in a homogeneous substance, i.e., it would have become unstable [S. V. Kozlovskaya (Ref. 18)]. One can assume that a small concentration of heavier substances to the center, created by differentiation led to the nearly homogeneous density of all the lunar interior or, if there is an iron nucleus in it, to an almost homogeneous density of the cover (mantel) of the Moon.

REFERENCES

1. N. A. Amenitskiy, R. I. Noskova, A. Ye. Salomonovich. Astromicheskiy Zhurnal, 37, 185-186, 1960.

2. D. Artur. Doklad na Simpoziume MAS (Report of the Symposium of the International Astronomical Association), No. 14, "Luna", (The Moon), 1960.

3. N. P. Barabashov and V. I. Yezerskiy. Izvestiya Komiteta po Fizike Planet (Report of the Committee on Planetary Physics), 2, 65-72, 1960.

4. N. P. Barabashov and V. I. Yezerskiy. Iav. Kom. po Fiz. Planet. (Report of Committee on Planetary Physics), 1, 67-79, 1959.

5. T. Weimer, Doklad na Simpoziume MAS, (Report of the IAU Symposium), No. 14, "Luna" (The Moon), 1960.

6. A. Wegener. The Origin of the Moon and of Its Craters (Proiskhozhdeniye Luny i Yeyo Kraterov) Gosudarstvennoye Izdatel'stvo, 1923.

7. A. P. Vinogradov. Izvestiya A. N., SSSR, Seriya Geologicheskaya, No. 10, 5-27, 1959.

8. A. P. Vinogradov. Khimicheskaya Evolyutsiya Zemli (The Chemical Evolution of the Earth) Izdatel'stvo A. N. SSSR, 1959.

9. T. Gold. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna" (The Moon) 1960.

10. J. Green. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna", 1960.

11. G. Jeffreys. Zemlya (The Earth) Izdatel'stvo Inostrannoy Literatury, 1960, p. 206.

12. Sh. Sh. Dolginov. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna" 1960.

13. A. Dollfus. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna" 1960.

14. M. R. Zelinskaya, V. S. Troitskiy, L. I. Fedoseyeva. Astronomicheskiy Zhurnal, 31, 643-647, 1959.

REFERENCES (Cont'd)

15. N. L. Kaidanovskiy, M. T. Turusbekov, and S. E. Kaikin. Trudy V Sovvyeshchaniya po Voprosam Kosmogonii (Proceedings of V Conference on Problems of Cosmogony) 1956, p. 347.

16. A. A. Kalinyak. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna", 1960.

17. J. Kuiper. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna", 1960.

18. S. V. Kozlovskaya. O Vnutrennyem Stroyenii Luny (On the Problem of the Moon's Inner Structure), This conference. p. 170.

19. N. A. Kozyrev. Izvestiya Krymskoy Astrofizicheskoy Observatorii, 16, 148, 1957.

20. N. A. Kozyrev. "Priroda", No. 3, 1959.

21. Z. Kopal. Doklad na Simpoziume MAS (Report of the IAU Symposium) No. 14, "Luna", 1960.

22. B. Yu. Levin. Voprosy Kosmogonii (Problems of Cosmogony), 6, 56-62, 1958.

23. B. Yu. Levin. Byulleten' Vsesoyuznogo Astronomo-geodezicheskogo Obshchestva (in printing) 1962.

24. B. Yu. Levin and S. V. Mayeva. Doklady A. N. SSSR, 33, 44-47, 1960.

25. A. V. Markov. Astron. Zh., 4, 60, 1927.

26. A. V. Markov. Astron. Zh., 25, 172, 1948.

27. Coll. Moon (Luna) A. V. Markov, ed., Moscow, Fizmatgiz 1960.

28. A. V. Markov. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna", 1960.

29. N. N. Pariyskiy. Astron. Zh. 37, 543-549, 1960.

30. J. Pettengill. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna", 1960. /143

REFERENCES (Cont'd)

31. H. I. Potter. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna", 1960.

32. V. V. Radzievskiy. Byul. Vses. Astron-Geodez Obshch., No. 11 (18), 3-8, 1952.

33. L. N. Radlova. Sc., Uchenyye Zapiski Leningradskogo Universiteta, 82, 99, 1941.

34. L. N. Radlova. Astron. Zh., 20, No. 1, 1943.

35. Ye. L. Ruskol. Astron. Zh., 37, 690-701, 1960.

36. Ye. L. Ruskol. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna", 1960.

37. A. Ye. Salomonovich. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna", 1960.

38. V. Sinton. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna", 1960.

39. N. N. Sytinskaya. Trudy Simferopol'skoy Astronomicheskoy Observatorii, 20, 106, 1954.

40. N. N. Sytinskaya. Uch. Zap. Leningrad. Univ. Ser. Mat. Nauk, Issue 29, No. 190, 1957.

41. N. N. Sytinskaya. Voprosy Kosmogonii (Problems of Cosmogony), 5, 13, 1957.

42. N. N. Sytinskaya. Astron. Zh., 36, 315, 1959.

43. F. Faut. Priroda Luny (The Nature of the Moon) St. Petersburg, 1911.

44. V. A. Fedorets. Uchennyye Zapiski Kharkovskogo Gosudarstvennogo Universiteta, 42, 49, 1952.

45. J. Frants. "Luna" (The Moon), Gosizdat 1923.

46. A. V. Khabakov. Ob Osnovnykh Voprosakh Istorii Razvitiya Poverkhnosti Luny (On Basic Problems in the History of Lunar Surface Development), Geograficheskoye Izdatel'stvo, 1949.

REFERENCES (Cont'd)

47. A. V. Khabakov. "Luna" A. V. Markov, ed., 241-297, 1960.

48. V. V. Sharonov. Trudy Astronomicheskoy Observatorii LGU (Proceedings of Astronomical Observatory of Leningrad State University), 6, 1936.

49. V. V. Sharonov. Uch. Zap. Leningrad Univ., 31, 28, 1939.

50. V. V. Sharonov. Astron. Zh., 31, No. 5, 442-452, 1954.

51. O. J. Schmidt. Chetyre Lektsii o Teorii Proiskhozhdenia Zemli (Four Lectures Concerning the Theory of the Origin of the Earth) Izd-vo A. N. SSSR, 1950.

52. H. Urey. Doklad na Simpoziume MAS (Report of the IAU Symposium), No. 14, "Luna", 1960.

53. A. A. Yakovkin. Byul. Astr. Obs. Imeni Engelgardta, No. 4, 1934.

54. A. A. Yakovkin. Izv. Astr. Obs. Imeni Engelgardta, No. 29, 1955.

55. D. W. C. Arthur. Journal of the British Astronomical Association, 64, No. 3, 1954.

56. J. Ashbrook. Publications of the Astronomical Society of the Pacific, 72, 55-58, 1960.

57. R. B. Baldwin. The Face of the Moon, Universiyt of Chicago Press, 1949.

58. R. B. Baldwin. Sky and Telescope, 21, No. 2, 84-85, 1961.

59. N. P. Barabashev. Astronomische Nachrichten, 221, 289, 1924.

60. E. Cartan. Proceedings of the International Mathematics Congress, Toronto, 2, 9-21, 1924.

61. J. van Diggelen. Bull. Astronom. Inst. Netherlands, 11, No. 423, 283-289, 1951.

62. J. van Diggelen. Rech. Astr. de l'Observ. d'Utrecht, 14, 2, 1959.

63. A. Dollfus. Annales d'Astrophysique, 19, No. 2, 1956.

64. J. Dubois. Comptes Rendus de la Societe de Physique de France, Oct., 1956.

REFERENCES (Cont'd)

65. E. Du-Fresne. Astronomical Journal 124, No. 3, 638-642, 1956.

66. H. Ebert. Astron. Nachr. 122, 263, 1889.

67. J. V. Evans. Proceedings of the Physical Society, London, 70, 1105, 1957.

68. B. Elsemore. Philosophical Magazine 2, No. 20, 1040, 1957.

69. J. Franz. Astronomische Beobachtungen der Sternwarte. Koenigsberg, 38, No. 5, 1899.

70. S. Fujinami, et al. Publications of the Astronomical Society. Japan. 6, 67, 1954.

71. T. A. Gehrels. Astronomical Journal, 64, 332, 1959.

72. L. A. Giamboni. Astrophysical Journal, 130, No. 1, 324-335, 1959.

73. T. M. N. Gold. Monthly Notices of the Astronomical Society 115, No. 6, 585-604, 1955.

74. T. Gold. Vistas in Astronautics, 2, 261, 1959.

75. C. R. Grant and B. S. Yaplee. Proceedings of Institute of Radio Engineers 49, 976, 1957.

76. J. Green. Proceedings of the Lunar and Planetary Exploration Colloquium, Griffith Observatory, 1-18, 1959.

77. J. H. Harring and L. Licht. Science, 130, 266, 1959.

78. J. H. Harring and L. Licht. Symposium of COSPAR, Nice, 1960.

79. V. A. Hughes. Nature, 186, No. 4728, 873-874, 1960.

80. Japlee, et al. Proc. I. R. E., 46, 293, 1958.

81. C. Jaschek. O. R. Observatory, 80, No. 916, 119-120, 1960. /144

82. H. Jeffreys. Monthly Notices Roy. Astron. Soc., 91, 169-173, 1930.

83. J. Joung. Journal of the British Astronomical Association, 50, No. 9, 309-326, 1940.

84. Z. Kopal. Nature 183, 169-170, 1959.

85. G. P. Kuiper. Proceedings of the National Academy of Sciences, Washington, 40, No. 12, 1096-1111, 1954.

REFERENCES (Cont'd)

86. G. P. Kuiper. Vistas in Astronautics, 2, 273, 1959. See also the Russian translation in the publication "Eksperimental'noye Issledovaniye Okolozemnogo Kosmicheskogo Prostranstva" (Experimental Research on Cosmic Space Near the Earth). Moscow, Izdatel'stvo Inostrannoy Literatury, 1961.

87. G. P. Kuiper. The Photographic Lunar Atlas, University of Chicago Press, 1960.

88. R. A. Littleton. Monthly Notices Roy. Astron. Soc., 98, 646, 1938.

89. B. Lyot. Annales de l'Observatoire de Paris, Sect. de Meudon, 8, 1929.

90. G. J. F. MacDonald. Planetary and Space Science, 2, No. 4, 249-255, 1960.

91. G. J. F. MacDonald. Science, 133, No. 3458, 1045-1050, 1961.

92. R. R. MacMath, et al. Publ. of the Observatory of Michigan 6, No. 8, 67-76, 1937.

93. N. Mannaert. Astrophys. J., 93, 403, 1941.

94. P. Moore. J. Brit. Astron. Assoc., 64, No. 1, 1954.

95. P. Moore. Sky and Telescope, 15, No. 5, 201, 1956.

96. J. Öhman. Stockholm Observatory Annales, 18, No. 8, 1950.

97. E. J. Öpik. Irish Astronomical Journal, 5, No. 1, 34-36, 1958.

98. E. J. Öpik. Astrophys. J., 132, No. 2, 502-503, 1960.

99. E. J. Öpik and Singer. S. F. Journal of Geophysical Research, 65, No. 10, 3057-3064, 1960.

100. G. H. Pettengill. Proc. I. R. E., 48, (5), 933-934, 1960.

101. E. Pettit. Astrophys. J., 91, 408, 1940.

102. E. A. Pettit, S. Nicholson. Astrophys. J., 21, 102, 1930.

103. H. Ritter. Astron. Nachr., 252, 157, 1934.

REFERENCES (Cont'd)

104. J. M. Saari and R. W. Shorthill. Astron. J., 66, No. 2, 1961.

105. Schrutka-Rechtenstamm und Hopmann. Sitzungsberichte Abteilung 2, Wein, 167, H. 8-10, 283, 290, 1958.

106. C. Seeger, et al. Astrophys. J., 126, No. 3, 585, 1957.

107. E. Shoemaker. Reports of the International Geological Congress XXXI Session, Norden. 1960. part XVIII Copenhagen, 1960.

108. E. Shoemaker, et al. Preprint, 7th annual meeting American Astronautical Society, 1961.

109. W. Sinton. Journal of the Optical Society of America, 45, 975, 1955.

110. W. Sinton. Lowell Observatory Bulletin, 5, 1, 1960.

111. Smoluchowski. International Critical Tables, 2, 3, 15, 1911.

112. J. H. Trexler. Proc. I. R. E., 46, 286, 1958.

113. H. Urey. Handbuch der Physik, Springer-Verlag, 3, 1959.

114. H. Urey, et al. Astrophys. J., 129, No. 3, 842-848, 1959.

115. K. Watson, B. Murray, and H. Brown. J. Geophys. Res., 66, No. 9, 3033-3045, 1961.

116. A. J. Wesselink. Bull. Astron. Inst. Neth., 10, 390, 352-361, 1948.

117. F. L. Whipple. Vistas in Astronautics, 2, 261, 1959.

ON THE PROBLEM OF THE MOON'S INNER STRUCTURE /145

By

S. V. Kozlovskaya

A number of models, made of ordinary silicates which are characteristic for the Earth and meteorites, are calculated. These models are divided into two groups: homogeneous models and those with a lighter crust, constituting 5, 10, and 15% by their mass, for two values of compressibility. The decrease of the Moon's density as a result of heating up of the interior and the decrease of the values of compressibility and the coefficient of thermal expansion are evaluated. Under these conditions, we have either a stable or an unstable model. The inner structure of the Moon is approximated by the two-layer model, with a shell that is inhomogeneous in composition but constant in density.

An attempt is made in this paper, to find out which deductions can be drawn concerning the inner structure of the Moon, on the basis of present data on the properties of rock formations and on the thermic history of the Moon.

Much data is available on the compressibility and the coefficient of thermal expansion of various minerals and rock formations for pressures up to 50,000 atmospheres, which are predominant in the Moon's interior. It is therefore possible to conduct direct calculations of lunar models, made of ordinary silicates. It follows from cosmogonical considerations that meteorites are approximately indicative of the matter content of the Earth and Moon. The composition of iron in chondrites amounts to about 10%. We did not consider this when we calculated the silicate models of the Moon.

As in the preceding works, (Refs. 1, 2) the models of the Moon are calculated by means of the numerical integration of the equation of hydrostatic equilibrium. The lunar mass used in these calculations is accepted as being equal to $7.32 \cdot 10^{25}$ gm, its radius is 1738 km, and the mean density 3.33 gm/cm^3.

Isothermic models of the Moon. A one-layer model. In the capacity of a first approximation we will calculate a fully homogeneous Moon model, for two values of compressibility: $\beta_1 = 1 \cdot 10^{-12}$ cm^2/dyne and $\beta_2 = 0.5 \cdot 10^{-12}$ cm^2/dyne. Within these limits lies the compressibility of those

minerals and rock formations (Ref. 3) of which the Moon should be primarily composed, if the assumptions that the Moon, Earth, and meteorites have an identical composition are correct. /146

The compressibility changes insignificantly within the interval of pressures up to 50,000 atmospheres (Ref. 3). Therefore, by disregarding this relationship, we can calculate the density according to the following formula

$$\rho = \rho_0 \ (1 + \beta P), \tag{1}$$

where ρ_0 is the density at a normal temperature and pressure, β is the compressibility, and P is the pressure.

The calculation has shown that in the case of β_1 it is possible to construct a lunar model, homogeneous in contents from a substance with a density of $\rho_0 = 3.26$ g/cm^3 and in the case of β_2 from a substance with $\rho_0 = 3.30$ g/cm^3 (Table 1).

<u>The two-layer model of the Moon</u>. A single-layer model of the Moon is, of course, far from the truth. The investigations on the thermal history of the Moon (Refs. 4, 5) shows that the interior of the Moon passed through a stage of partial or even full melting. On the surface of the Moon, we see traces of enormous lava flows. The melting must have caused a stratification of the Moon and the formation of a crust. This should be taken into consideration during the calculation of the model.

However, we do not know what portion of the Moon's mass its crust contributes. Due to the variation between the masses of the Earth and Moon and the pressures in their interiors, the process of matter differentiation by density and chemical composition, i.e., the processes which lead to the formation of a crust, have occurred somewhat differently on the Earth and on the Moon. In the case of the Earth, the differentiation is still far from complete and thus, the mass of the Earth's crust is very small, less than 1%. Calculations show that from meteoric type matter, a maximum of about 7% SiO_2 (Refs. 6, 7) can be extracted. At a full extraction, the Earth's crust might reach 15% by mass. /147

At low pressures in the interior of the Moon, the melting and the escape of lighter substances toward the surface begins at a significantly lower temperature than the temperature in the interior of the Earth. Thereby, the escape of light substances is facilitated by the fact that the melting on the Moon signifies a transition of matter into

Table 1. One-Layer and Two-Layer Models of the Moon Made of Substance with a Constant Compressibility

Mass of the crust, %	Thickness of the crust, km	β (compressibility) of the substance of the Moon, $cm^2/dyne$	P (pressure) at the base of the crust in atmospheres	ρ_0 (density at STP) of the substance below the crust, gm/cm^3	ρ (density) of the shell at the base of the crust gm/cm^3	ρ (density) in the center, gm/cm^3	P_c pressure in the center, in atmospheres
a) Isothermic Models							
0	0	$1 \cdot 10^{-12}$	-	3.26	-	3.43	50 000
0	0	$0.5 \cdot 10^{-12}$	-	3.30	-	3.38	50 100
5	35	$1 \cdot 10^{-12}$	1570	3.30	3.305	3.47	51 000
5	35	$0.5 \cdot 10^{-12}$	1570	3.325	3.326	3.41	52 000
10	72	$1 \cdot 10^{-12}$	3220	3.33	3.341	3.50	50 300
10	72	$0.5 \cdot 10^{-12}$	3220	3.365	3.370	3.46	50 700
15	110	$1 \cdot 10^{-12}$	4930	3.37	3.386	3.55	52 500
b) Non-Isothermic Models, $\alpha = 3 \cdot 10^{-5}$ deg^{-1}							
0	0	$1 \cdot 10^{-12}$		3.345		3.375	52 000

a liquid volatile state. Therefore, with a sufficiently great content of radioactive elements, it is possible to assume an almost complete differentiation. Nevertheless, it is possible that such a division was not general. Thus, according to Kuiper (Ref. 8) the bright continents are actually ancient crust deposits, i.e., sections of the Moon's non-differentiated substance.

The considerations indicated above do not permit us to evaluate as yet the potential of the Moon's crust. Therefore, we have calculated some two-layer models of the Moon for three versions of crust thickness, with the assumption that the lunar crust, having a density of $\rho_0 = 2.8$ gm/cm^3 (exactly as in the Earth's crust), amounts to 5, 10 and 15% by mass. In any case, the mass of the actual crust of the Moon should not exceed the indicated upper limit. The models were calculated from matter with a constant compressibility identical by crust and shell for two values of β. The calculation has shown that in the two-layer silicate models of the Moon, the density of the shell substance can be quite high. Immediately beneath the crust, the density is close to or exceeds the mean density of the Moon (Table 1).

Lunar models with the consideration of a temperature change along the radius. We will however, examine the influence which the consideration of temperature has on the calculation results. The dependence of density on temperature t is determined by the formula

$$\rho = \frac{\rho_0}{1 + \alpha t}, \tag{2}$$

where α and β according to formulae (1) and (2) are determined by the following expression,

$$\rho = \rho_0 \frac{1 + \beta P}{1 + \alpha t}. \tag{3}$$

The results of the calculation of the present temperature distribution in the Moon's interior (Ref. 5) depends essentially on the /148 accepted concentration of radio-active elements, on their distribution, and on heat conductivity. In this work, we have found the temperature distribution along the Moon's radius corresponds to the average values of the parameters which are examined in the work (Ref. 5). The progress of pressure and temperature along the Moon's radius is shown in Table 2 (R_0 is the radius of the Moon, r is the distance from its center).

Table 2. Distribution of Pressure and Temperature Along the Radius of the Moon Adopted for Calculation

r/R_0	1.0	0.8	0.6	0.4	0.2	0
t,°C	0	800	1 185	1 360	1 420	1 450
P, atm.	0	16 600	30 000	39 600	46 100	51 000

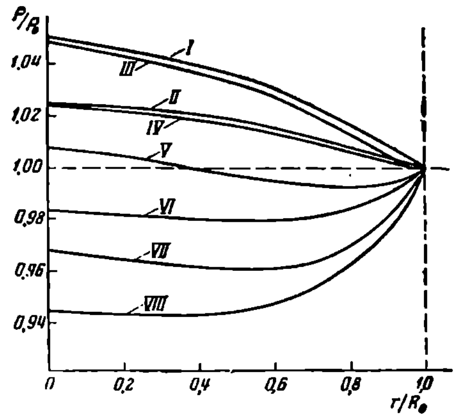
Strictly speaking, it is necessary to conduct the calculations by way of consecutive approximations since the distribution of pressure is related to the distribution of density. However, the pressure distribution is very stable, as far as various suppositions on the inner structure are concerned, as can be seen in Table 1, where the pressures in the centers of various models are shown (the pressure variations are at a maximum in the center). As a comparison, calculation data for a non-isothermic model is given. Therefore, in Table 2, the mean pressure distribution is taken graphically from the pressure distribution of the models in Table 1.

The non-isothermic models of the Moon are calculated for the same values of β and for two values of α: $\alpha_1 = 3 \cdot 10^{-5}$ and $\alpha_2 = 6 \cdot 10^{-5}$ deg^{-1}. The value of α obtained experimentally for olivine (Ref. 3) approximates α_1 and the theoretical α for the substance of the Earth (Refs. 9, 10) approximates α_2. From formula (3) and Table 2, the change in the magnitude ρ/ρ_0 along the radius of the Moon for accepted values of α and β and for the accepted temperature distribution in its interior was determined.

In the upper parts of the Moon where the pressure is still small, the influence of temperature is so significant that the density decreases with depth. When $\beta < 1 \cdot 10^{-12}$ cm^2/dyne and $\alpha > 3 \cdot 10^{-5}$ deg^{-1}, the inversion of densities is found at all depths. The stable model in which the density increases with depth and the layers are chemically homogeneous is found with the following values for α and β: $\beta > 1 \cdot 10^{-12}$ cm^2/dyne and $\alpha < 3 \cdot 10^{-5}$ deg^{-1} (see illustration).

The actual picture is, of course, quite different from the picture seen in the illustration. The formation of the crust led to a decrease

in mean density in the surface layer. On the other hand, with the evolution of heat and softening of the lunar interior, the denser /149 components should have sunk, and the lighter should have floated up eliminating the inversion of densities. As a result of the differentiation by chemical composition, the density ρ_0 of the substance of the Moon should increase from the surface to the center.



Change in ρ/ρ_0 along the Radius R_0 of the Moon (r is the distance from the center of the Moon) for various values of β and α.

Isothermic models: I and II are one-layer models for $\beta = 1 \cdot 10^{-12}$ and $0.5 \cdot 10^{-12}$, respectively; III and IV are two-layer models, the curst is 10% in mass, $\beta = 1 \cdot 10^{-12}$ and $0.5 \cdot 10^{-12}$, respectively.

Non-isothermic models: V - $\beta = 1 \cdot 10^{-12}$, $\alpha = 3 \cdot 10^{-5}$ deg^{-1}; VI - $\beta = 0.5 \cdot 10^{-12}$; $\alpha = 3 \cdot 10^{-5}$; VII - $\beta = 1 \cdot 10^{-12}$, $\alpha = 6 \cdot 10^{-5}$; VIII - $\beta = 0.5 \cdot 10^{-12}$, $\alpha = 6 \cdot 10^{-5}$.

The effect of temperature on the density of the substance in the lunar interior is greater than the action of pressure. Nevertheless, the increase of density due to the change in chemical composition with depth could have led to the fact that the density in the Moon's interior could have proved to be almost constant.

Thus, one can construct a model of the Moon from the same rock formations which, as assumed, are the most common in the composition of the Earth. Since the density change in the Moon's interior is small, then one can approximate the inner structure of the Moon by means of a two-layer model in which there is a crust having a constant density and a core that is homogeneous in chemical composition with a different but still constant density.

REFERENCES

1. S. V. Kozlovskaya. Doklady Akademii Nauk SSSR 92, No. 5, 903-906, 1953.

2. S. V. Kozlovskaya. Doklady A. N. SSSR 108, No. 3, 409-412, 1956.

3. F. Birch and others. Spravochnik dlya Geologov po Fizicheskim Konstantam (Handbook of Physical Constants for Geologists) Izdatel'stvo Inostrannoy Literatury, 1949.

4. B. Yu. Levin and Ye A. Lyubimova. "Priroda", No. 10, 81-84, 1955.

5. B. Yu. Levin and S. V. Mayeva. Doklady A. N. SSSR, 133, No. 1, 44-47, 1960.

6. A. P. Vinogradov. Izvestiya A. N. SSSR, Seriya Geologicheskaya, No. 10, 5-27, 1959.

7. A. P. Vinogradov. Khimicheskaya Evolyutsia Zemli (The Chemical Evolution of the Earth) Izdatel'stvo A. N. SSSR, 1959.

8. G. P. Kuiper. Proceedings of the National Academy of Sciences of the USA, 40, No. 12, 1096-1112, 1954.

9. V. N. Zharkov. Izv. A.N.SSSR, Ser.Geofiz., No. 9, 1414-1419, 1959.

10. J. A. Jacobs. Handbuch der Physik, 47, 364-406, Berlin, Springer-Verlag, 1956.

ON THE PROBLEM OF THE ROTATION OF PLANETS /150

By

V. S. Safronov

In this paper is analyzed the explanation of the rotation of planets, which was suggested by O. J. Schmidt, based on his theory that planets were formed through a junction of solid particles and bodies of a protoplanetary cluster. An analysis of the equation for energy and momentum balance of the amount of motion, revealed the inadequacy of the simplified system which O. J. Schmidt considered, i.e., the system which suggests that the particles of the cluster traveled along circular orbits. The condition of direct rotation which O. J. Schmidt found is not fulfilled in this system at the present rotation speed of the Earth. It would have been necessary for the Earth to rotate 10^4 times faster, in order to fulfill this condition.

The angular and energy momentum equations for a more general case of the motion of bodies in elliptical orbits are deduced. It is shown that in this new model, the contradictions mentioned above are eliminated.

The relation between the planet rotation and the heat losses during the formation process is analyzed. A revision of the former concept of that relation is dictated by the consideration of the planet's rotation energy. It has been shown that the heat losses diminish with the increase of speed of rotation.

It follows from the expressions obtained that the simplest supposition about the velocity of rotation derived from dimensional considerations leads to the proportionality between the rotational energy increase and the potential energy on the surface of the growing planet. This in turn leads to the approximate constancy of angular velocity during the process of growth, in agreement with the results obtained earlier on the assumption of the asymmetry of shocks of falling bodies and with the fact that the differences in the periods of rotation of the planets are comparatively small.

In Laplace cosmogony, the direct rotation of the planets is /151 usually connected with the action of the Sun's tidal forces. Poincaré (Ref. 1) gives the following schematic description of this process. Due to the forces of friction, the gaseous ring begins to rotate around the Sun as a rigid body. Subsequently, it becomes unstable and breaks apart. The individual parts of the ring begin to move along circular orbits.

The joining of two clouds, of which the distances from the Sun differ slightly leads to their opposite rotation since the inner cloud orbits faster than the outer. However, the tidal forces of the Sun stretch this cloud out and impart to it a direct rotation with a period equal to its period of revolution. The contraction of the clouds due to their cooling, decreases the effect of tidal forces and increases their velocity of rotation.

At present none of these suppositions are acceptable. The forces of friction in a ring rotating around the Sun do not cause it to rotate like a rigid body. They merely draw the inner part of the ring nearer to the Sun and move out the outer parts, the rotation all the while remaining Keplerian. The idea of the break-up of the protoplanetary cloud into small gaseous clouds and their subsequent rejoining is also unacceptable. Such clouds are unstable and would tend to disintegrate farther, not to rejoin. This difficulty could be avoided only by dropping the gaseous cloud concept in favor of that of dust concentration, but dust concentrations are small and shrink so rapidly that the tidal forces could not perceptibly change their rotation (Ref. 2).

V. G. Fesenkov (Ref. 3) explains the rotation of planets using the hypotheses of massive protoplanets. It is assumed that the density of the gaseous component of the cloud had reached the critical value (Roche density) and that the set-in instability had caused massive clots to form in the cloud. The process of condensation is conceived as an accumulation of matter along the initial clot's orbit. It is further assumed that turbulent mixing had led to the equalization of angular velocity of rotation in the cloud. For the determination of planet rotation, the moment of the momentum of the cloud rotating as a rigid tore with a diameter of 2ℓ around the axis through the center of its cross section is calculated. Then, ℓ is chosen such as to make the moment of that tore equal to the rotating moment of the initial planet rotating with an angular velocity of rotation equal to the present one. The moment of the tore emerges so great that we end up with a very small value of ℓ. The /152 diameter of a tore having the rotating moment of Jupiter was found to be equal to 59 diameters of the planet, i.e., 70 times smaller than the width of Jupiter's zone. For the protoplanet Earth having a mass 30 times greater than at present, the diameter of the tore would be 300 times the width of the Earth's zone. With the gravitational instability in a flat, rotating layer, the dimensions of the non-disintegrating clots should exceed the thickness of the layer by one order (Refs. 4, 5). The previously indicated values of ℓ do not satisfy this condition - they are a magnitude of 1.5 orders smaller in the Earth's zone and of 1 order in Jupiter's zone. Hence, the protoplanets could not be formed by the consolidation of substance included in a tore of such a small cross section. But the real non-disintegrating clots would have had a specific moment bigger by two orders than the necessary moment in Jupiter's zone and by three - in the Earth's zone.

The turbulence in the rotating medium equalizes not the angular velocity but rather the moment of the momentum (Ref. 6); besides, the turbulence in the protoplanetary cloud should have been dying off. There is therefore no basis for the assumption that the tore is rotating as a rigid body and it would be more natural to proceed from Kepler's rotation. It is necessary to note that in this case, the rotation of the forming consolidations is also found to be direct, except that the velocity proves to be somewhat smaller (Refs. 2, 7). Indeed, the mean velocity of rotation of a local area having a circular form is determined by the equation

$$\bar{\omega} = \left|\frac{1}{2} \text{rot } V\right| = \frac{1}{2R} \frac{d}{dR} (V R),$$

where V is the linear velocity at a distance R from the axis of rotation. With the Kepler circular motion $V = \sqrt{GM_{\odot}/R}$ and $\bar{\omega} = (1/4)\ \omega_c$, where $\omega_c = \sqrt{GM_{\odot}/R^3}$. Thus, at the Keplerian motion of the substance, the moment of the momentum of the formed protoplanet is merely four times smaller than at a rigid rotation, but still too great.

The protoplanets hypothesis meets with no less serious difficulties with regard to the mass of the protoplanetary cloud. For the gravitational instability in the gas, it is necessary for the cloud's mass to be not less than the Sun's mass (Ref. 8). At Roche's density in the Earth's zone alone, the mass ΔM_3 of the gas of the Earth's zone alone should reach one tenth of the Sun's mass. This is directly observed from the relationships $\Delta M_3 \approx 2\pi R\rho_R\ H \Delta R$; $H \approx 10^{-2}\ R$, $\Delta R \approx \frac{1}{2} R$ and according to V. G. Fesenkov, $\rho_R = 4M_{\odot}/R^3$. Here at once emerge two problems: 1) Why did merely one thousandth of the entire zone's mass go into the protoplanet formed in the Earth's zone? 2) What forces could have removed /153 from the solar system a mass of the order of the Sun's? J. Kuiper (Ref. 9) showed that it would be impossible to explain the dissipation from the solar system of a mass greater than 1/10th the mass of the Sun. The difficulties with the dissipation and sorting out of light gases from the massive protoplanets were especially emphasized by I. S. Shklovsky (Ref. 10). The first question also remains unexplained.

The excessively great moment of rotation of the planets which we find by assuming their formation from massive protoplanets which have condensed due to gravitational instability in the gaseous component of the protoplanetary cloud is one more item against this hypotheses.

F. Hoyle (Ref. 11) proposed an explanation of the planets' rotation proceeding from the conception of the planet's growth through the accretion by the massive nucleus of a planet of the scattered substance visualized as a continuous medium. As the probable radius of capture for the massive nuclei (mass m and radius r) he assumed half the distance at which the tidal force of the Sun is equal to the gravity pull of the nucleus: $r_a = 1/2\ (m/2M_\odot)^{1/3}\ R$. The rotational moment ΔK which is brought in by the accrueing substance Δm is taken $2/5\ \overline{\omega} r^2 \Delta m$ (spherically symmetrical accretion), where, as above, $\overline{\omega} = 1/4\ \omega_c$. The period of rotation has proved to be equal to 3-4 hours without taking into account the concentration of substance toward the center of the planet, which would have involved still faster rotation. Hoyle allowed for such a great velocity of rotation under the influence of Littleton's idea of the rotational instability of the primary planets and the separation of satellites from them. Although Littleton has supported this idea up to the present, (Ref. 12), its bases are not sufficiently valid. On the other hand, "accretion" played a substantial role only during the growth of Jupiter and Saturn when their embryos became sufficiently massive for absorption of gaseous hydrogen. Recently even Hoyle (Ref. 13) himself abandoned the idea of applying the mechanism of accretion to the planets in the Earth's group. As for the rotation of Jupiter and of Saturn it apparently could be satisfactorily explained within the bounds of the theory of accretion. However, for this, a closer estimate of the "capture" radius would be necessary. If it turns out to be two or three times smaller than the value taken by Hoyle, it could be regarded as confirming the theory.

L. E. Gurevich and A. I. Lebedinskiy (Ref. 4) have shown the feasibility of the idea of the protoplanetal dust disc's breaking up into numerous concentrations, the consolidation of which should have led to the formation of planets. Having determined the masses and radii of the concentrations, they found that the rotation moment of the concentration was proportional to the expression $k_0 M_n (M_n/M_0)^4$, where k_0 is the unit orbital moment and M_n is the mass of the planet. From this it was deduced that the rotation moment of the planet should be equal to the orbital moment multiplied by a function of the planet's mass. The deduction /154 was illustrated by an empirical relation which checks out well for all planets except Saturn and Neptune. This conclusion is essentially founded on the non-self-evident supposition of central collisions of the consolidating concentrations. At non-central collisions of the consolidating concentrations - besides their own rotation moments, account must also be taken of the considerably larger moments connected with their relative orbital motion. Until such an analysis is made it would be imprudent to extend the relationship found for the rotation moments of the concentrations to planets.

In his analysis of the problem of planet rotation O. J. Schmidt (Ref. 14) proceeded from consideration of the general principles of the process of consolidation of matter into a planet. He wrote out the conditions for the conservation of energy and of the moment of momentum during the transition from a cloud of particles to a system of planets. The full moment of the particles located in a planet zone is transformed into the orbital and rotation moments of the planet. The smaller the orbital moment, i.e., the smaller the radius of the planet's orbit, the greater the rotation moment must be. But the smaller the radius of the planet's orbit, the smaller is its orbital energy, and hence, according to O. J. Schmidt, the greater are the thermal energy losses in the process of the planet's formation. From this O. J. Schmidt draws his main deduction: since the losses of energy in this process are great, then the planet's rotation must be direct. The mathematical formulation of the result boils down to the following.

A cloud of particles moving in a plane circular orbit around the Sun is being considered. From these particles a planet with an orbit radius of R_0 is formed. Let R_e and R_m be the mean distances of the particles in the cloud, derived for the energy and the moment respectively:

$$\frac{1}{R_e} = \frac{\int_{R_1}^{R_2} \frac{\varnothing\,(R)}{R}\, dR}{\int_{R_1}^{R_2} \varnothing\,(R)\, dR}, \quad \sqrt{R_m} \approx \frac{\int_{R_1}^{R_2} \sqrt{R}\,\varnothing\,(R)\, dR}{\int_{R_1}^{R_2} \varnothing\,(R)\, dR}, \tag{1}$$

where $\varnothing\,(R)$ is the function of the particle mass distribution according to their distances from the Sun, and R_1 and R_2 are those of the boundary of the zone of the subject planet.

It can be shown that R_m is always greater than R_e. Therefore, if the condition

$$R_0 < R_e, \tag{2}$$

is satisfied then the inequality /155

$$R_0 < R_m, \tag{3}$$

should be satisfied, i.e., the rotation should be direct, since the orbital moment of the planet which is proportional to $\sqrt{R_0}$, is less than the moment of the cloud which is proportional to $\sqrt{R_m}$.

No analysis was made of the conditions at which relationship (2) is satisfied. O. J. Schmidt assumed that "we cannot determine the sum of these losses quantitatively, but there is no doubt that the losses are great". He further assumed that the same causes condition the direct revolution of a majority of the planet's satellites and that the opposite revolution of distant satellites is connected with the failure to fulfill the conditions (2). It is apparent from this that the problem concerning the planets' rotation requires a more detailed quantitative analysis.

Let us examine again the balance equations of energy and momentum moment derived by O. J. Schmidt. In the case of particles moving along circular orbits the sum of their kinetic and potential energy relative to the Sun is equal to

$$- \frac{GM}{2} \int_{R_1}^{R_2} \frac{\emptyset\ (R)}{H}\, dR,$$

While their total momentum moment relative to the Sun is equal to

$$\sqrt{GM} \int_{R_1}^{R_2} \sqrt{R}\, \emptyset\ (R)\, dR,$$

where M is the mass of the Sun, and G is the gravity constant. Let us introduce the symbols:

U_0 - the potential energy of the planet relative to the Sun,

U_p - the potential energy the planet as a sphere,

U_c - the potential energy of the particles' interaction with one another,

E_0 - the orbital energy of the planet (the potential plus the kinetic),

E_r - the kinetic energy of the planet's rotation,

E_t - the losses of energy of the mechanical motion due to transformation into different forms of energy -- heating, radiation, phase transitions, etc.,

m - the mass of the planet,

R_0 - the radius of the planet's orbit,

K_0 - the planet's orbital moment,

K_r - the planet's rotational moment.

The balance equations will appear in the form /156

$$\left.\begin{aligned} -\frac{GM}{2}\int_{R_1}^{R_2}\frac{\varnothing(R)}{R}\,dR + U_c &= E_0 + U_p + E_r + E_t, \\ \sqrt{GM}\int_{R_1}^{R_2}\sqrt{R}\,\varnothing(R)\,dR &= K_0 + K_r. \end{aligned}\right\} \tag{4}$$

It is obvious that

$$E_0 = -\frac{GMm}{2R_0}, \quad K_0 = m\sqrt{GMR_0}, \quad m = \int_{R_1}^{R_2}\varnothing(R)\,dR. \tag{5}$$

The relationships (4) can be rewritten in the form

$$\left.\begin{aligned} \int_{R_1}^{R_2}\frac{\varnothing(R)}{R}\,dR - \frac{1}{R_0}\int_{R_1}^{R_2}\varnothing(R)\,dR &= -\frac{2}{GM}(U_p + E_t + E_r - U_c), \\ \int_{R_1}^{R_2}\sqrt{R}\,\varnothing(R)\,dR - \sqrt{R_0}\int_{R_1}^{R_2}\varnothing(R)\,dR &= \frac{K_r}{\sqrt{GM}}, \end{aligned}\right. \tag{6}$$

or, taking into consideration (1), in the form

$$R_0 = (1 - \epsilon)\, R_e,$$
$$R_m = (1 + k)^2\, R_0, \tag{7}$$

where

$$\epsilon = \epsilon_p + \epsilon_t + \epsilon_c = \frac{2R_0}{GMm}\,(U_p + E_t + E_r - U_c),$$
$$k = \frac{K_r}{m\sqrt{GMR_0}}. \tag{8}$$

eliminating R_0 we get

$$(1 + k)^2\,(1 - \epsilon) = \frac{R_m}{R_e}, \tag{9}$$

in view of the smallness of k and ϵ in comparison with the unit

$$2k - \epsilon \approx \frac{R_m - R_e}{R_e}. \tag{10}$$

The relationship $(R_m - R_e)/R_e$ is mainly dependent on the width of the zone supplying the planet and is only slightly dependent on $\varnothing(R)$. For the Earth's zone, it is $\sim 10^{-2}$ [see (18)]. The ratio of the rotation to the orbital moment of the Earth $k \approx 3 \cdot 10^{-7}$. With these numerical /157 values, it follows from the second equation (7):

$$R_0 \approx R_m\,(1 - 2k) > R_e. \tag{11}$$

Thus, the stipulation for direct rotation of planets (2) arrived at by O. J. Schmidt for the case of particles in circular orbits is not fulfilled for the Earth. If it were fulfilled, and $R_0 < R_e$ then the Earth would have rotated 10^4 times faster than it does now.

If we were to assume that condition (2) is not an indispensible but merely a sufficient condition for direct rotation of the planets and that the actual R_0 is determined from (11), another difficulty would

arise then: at k corresponding to the present rotation of the Earth, the relationship (10) is not fulfilled. The right side, as we have already seen, is $\sim 10^{-2}$. The components e are: $\epsilon_p \approx - 10^{-1}$, $\epsilon_r \approx 10^{-4}$, $\epsilon_c \approx 10^{-7}$, and $\epsilon_t \approx -\epsilon_p$ and it is hardly less than $- \epsilon_p - \epsilon_r$. Therefore, the left side appears to be of a magnitude of at least two orders smaller than the right. In order to satisfy (10) we would have to reduce the thermal losses to a magnitude of approximately $- 0.9\ \epsilon_p$. It is uncertain whether such an action is warranted.

One has to look for the cause of such a result in the shortcomings of the system itself. Generally speaking, the assumption of initial motion of the particles in circular orbits seems natural. At the small masses, of the planets' embryos, their gravitational disturbances were weak and the particles moved along orbits which were close to circular. With the growth of the embryo, and of other bodies, the deviations from circular orbits increased, and all the bodies in the zone were gradually consolidating into one planet. The above described balance would have been quite correct if the planet zone all the while remained self-contained. However, the result arrived at earlier indicates the inaccuracy of just such an assumption. As the eccentricities of the bodies increase due to their drawing together with the embryo and with one another, a part of them moves out beyond the outer boundary of the zone and remains there (sticks), carrying away from the zone a surplus moment, another part moves out beyond the inner boundary of the zone carrying off a moment which is below the mean moment. Simultaneously bodies from the outside are moving into the zone. These processes of interchange do not compensate each other, as a result of which the full moment of the momentum of the substance in the zone and the total energy (including thermal losses) do not remain constant.

The formulation of the problem would be more correct if one took into consideration the eccentricities and inclinations of the orbits of the bodies and particles and included in the function of their distribution along the major semi-axis Ø (a) only those which actually get /158 on to the embryo. Then the difficulties connected with the crossings of the zone's boundaries and with the increase in the eccentricities of the orbits and of the zone's dimensions, could be by-passed to a considerable degree by analyzing not the full balance between the initial and final states, but rather the "differential" balance, i.e., the balance for a set of given values of m, e and i and the one due to a small increment Δm of the embroy's mass. It would not be necessary to consider specifically the moment and energy exchange between the bodies getting on the embryo and those that do not, because the orbit eccentricities e and inclinations i, taken as characteristics of the initial state would be the very result of such exchange. One could, apparently,

completely disregard the effect on the motion of the embryo from the bodies which did not get on it, though the possibility could not be completely excluded that the embryo moving on its circular orbit might experience a slight braking effect as a result of coming close to bodies, whose centroid velocity is, as is known, somewhat smaller than the circular. However, the decrease of the embryo's orbital radius connected with it should, apparently, be quite insignificant.

For the sake of simplicity, we will assume that all bodies and particles have identical eccentricities and inclinations to the plane of the embryo's orbit. We will emphasize that by a, e and i we mean the average undisturbed elements, i.e., those corresponding to the time intervals when the body is not found in the state of consolidation with other bodies. Let us assume that while the substance quantity Δm is getting on the embryo, the bodies and particles having orbits whose large semi-axis are between a and a + da, bring in the mass $\Delta m \varnothing (a)\, da$. Then $\int_{a_1}^{a_2} \varnothing (a)\, da = 1$. Let us further assume that the embryo is moving in a circular orbit ($a_0 = R_0$). The potential energy of the interaction among particles and bodies outside of a consolidation is very small and can be excluded.

Instead of the equations (6) we get

$$\left.\begin{aligned} &\Delta m \int_{a_1}^{a_2} \frac{\varnothing (a)}{a}\, da - \Delta \left(\frac{m}{a_0} \right) \int_{a_1}^{a_2} \varnothing (a)\, da = - \frac{2}{GM} \Delta \left(U_p + E_t + E_r\right), \\ &\Delta m \int_{a_1}^{a_2} \sqrt{a\left(1 - e^2\right)} \cos i \varnothing (a)\, da - \Delta \left(m \sqrt{a_0}\right) \int_{a_1}^{a_2} \varnothing (a)\, da = \frac{\Delta K_r}{\sqrt{GM}} . \end{aligned}\right\} \quad (12)$$

We also have

$$\Delta \left(\frac{m}{a_0} \right) = (1 - \gamma) \frac{\Delta m}{a_0}, \quad \Delta \left(m \sqrt{a_0}\right) = \left(1 + \frac{\gamma}{2}\right) \sqrt{a_0} \Delta m,$$

$$\text{where } \gamma = \frac{d \ln a_0}{d \ln m} .$$

Applying expressions (1) to the major semi-axes of the orbits, /159 instead of to radii, we obtain from (12), by analogy with (7),

$$a_0 = (1 - \epsilon' - \gamma)\, a_e,$$

$$(1 - e^2)\cos^2 i\; a_m = \left(1 + k' + \frac{\gamma}{2}\right) a_0, \tag{13}$$

where

$$\epsilon' = \frac{2a_0}{GM}\frac{d}{dm}(U_p + E_t + E_r),\; k' = \frac{1}{\sqrt{GM\, a_0}}\frac{dK_r}{dm}. \tag{14}$$

Eliminating a_0 we will get

$$(1 - e^2)\cos^2 i\; a_m = (1 - \epsilon' - \gamma)\left(1 + k' + \frac{\gamma}{2}\right)^2 a_e. \tag{15}$$

The magnitude γ which characterizes the change in length of the orbital radius of the planet's embryo can be evaluated from the second equation (12) by making in it $\varnothing(a) = ca^{-n}$ and $\Delta K_r = 0$ because of its smallness. Calculations show that with the acceptable values of n the magnitude of γ will be of the order of e^2. In (15) γ enters only in the terms of e^4 and higher powers of e. Confining ourselves by small values to the second order for e and i and to the first order for k', we obtain

$$\frac{a_m - a_e}{a_e} \approx e^2 + i^2 + 2k' - \epsilon'. \tag{16}$$

Comparison of (16) and (10) reveals that the taking into account of eccentricities and inclinations of the orbits introduces additional terms e^2 and i^2 of the same order as the others which can substantially change the result. Hence, a system based on the analysis of circular orbits of particles is clearly inadequate.

For the determination of the left part of (16) we will introduce a dimensionless "distance"

$$x = \frac{a - a_1}{a_1}$$

and let us express the distribution function $\emptyset$ (a) in terms of x confining ourselves to terms of the second order

$$\emptyset = \emptyset_1 \ (1 + c_1 x + c_2 x^2). \tag{17}$$

Having substituted this expression for $\emptyset$ in (1), in which, instead of R we must now use a throughout, we can find a_m and a_e. The calculations show that, correct to the third power of x, the expression (16) is not dependent upon c_1 and c_2:

$$\frac{a_m - a_e}{a_e} = \frac{x_2^2}{16} \ (1 - x_2 + \cdot \cdot \cdot). \tag{18}$$

The value x_2 corresponds to the outer boundary of the zone a_2. /160
Taking the zone boundaries according to O. J. Schmidt or to L. E. Gurevich and A. I. Lebedinskiy (they differ but little) we obtain for the Earth $x_2 \approx 0.6$.

Since

$$a_1 = \frac{R_0}{1 + e} , \quad a_2 = \frac{R_0}{1 - e} , \tag{19}$$

then

$$x_2 = \frac{a_2 - a_1}{a_2} = \frac{2e}{1 - e} \approx 2e, \quad \frac{a_m - a_e}{a_e} \approx \frac{e^2}{4} .$$

Substituting this value in (16) we find

$$\epsilon' \approx \frac{3}{4} + i^2 + 2k', \tag{20}$$

or

$$E_t' + E_r' \approx - U_p' + \frac{1}{2} V_c^2 \left(\frac{3}{4} e^2 + i^2 \right) + k' V_c^2 . \tag{21}$$

This expression can be presented in a more obvious form by introducing the velocity v of the bodies relative to the embryo and not yet affected

by its pull (the velocity before the joining). The component of the velocity perpendicular to the plane of the embryo's orbit $v_r \approx iV_c$ where V_c is the circular velocity and the component in the plane of the orbit $v_{R\theta} \approx eV_c \sqrt{1 - \cos^2 \psi}$, where ψ is the angular distance of the body from perihelion at the moment of encounter. Let $v^2_{R\theta} = \lambda e^2 V^2_c$. Then

$$v^2 \approx (\lambda e^2 + i^2) V_c^2 \tag{22}$$

and from (21) we find

$$E_t' + E_r' \approx - U_p' + \frac{v^2}{2} + k'V_c^2 + \frac{1}{2}\left(\frac{3}{4} - \lambda\right) e^2 V_c^2. \tag{23}$$

The exact value of λ is difficult to calculate. For this, generally speaking, it is necessary to know the density distribution in the cluster $\emptyset$ (a). In any case λ is close to 3/4 and the last term of (23) is at least by one order smaller than V^2.

The examination of this equation leads to the following deductions.

1. At the accuracy to small e^2 with which the calculations were carried out, the thermal losses occurring at the landing of particles on the surface of the planet embryo are equal to the sum of the potential energy released by the particles at landing and of their kinetic energy prior to joining the embryo. (Precisely such a magnitude of losses was assumed by us (Ref. 15) in the evaluation of Earth's initial temperature). /161 Here we do not encounter the contradiction which emerged at the analysis of circular orbits, when the natural assumptions regarding the losses [condition (2)] led to inadmissibly high rotation values.

2. The relationship (23) derived from the balance equations contains two unknown quantities -- the velocity of the planet's rotation and the thermal losses during the process of accumulation. Hence, it is inadequate for the solution of the question regarding the rotation of the planets. Only the analysis of a concrete mechanism of collisions makes it possible to determine, with the help of balance equation, both the rotation and the losses.

3. From (23), at first glance, it appears that the greater E_t' is, the greater k' is, i.e., at great thermal losses the rotation should be direct. However, the actual picture is much more complicated. The term with k' is the smallest in (23) and it would be futile to attempt to

determine it from it, since it appears in (23) as a difference of two values greater than it by four orders. It is therefore impossible to tell from (23) anything concerning the direction of the planet's rotation. Since at the present velocity of the Earth's rotation the term k' is by 2.5 orders smaller than E_r' that characterizes the energy of the rotation, it is more correct to judge the connection between the rotation and losses not by k' and E_t' but rather by E_r' and E_t'. But these values appear everywhere as a sum. Hence, for given eccentricities of the bodies' orbits, the faster the rotation of the planet, the smaller should be the thermal losses of the process of accumulation. This deduction is physically conceivable from the following: The rotation is accelerated more intensively, (a) the greater the number of particles impinging in the direction of rotation, and (b) the amaller the number of particles impinging against the rotation, i.e., the smaller the velocity at which the impingements occur; and hence the smaller the thermal losses. By the way, this deduction is also valid for O. J. Schmidt's analysis of the motion of bodies in circular orbits. In his balance equations, the energy of the losses also appears together with the energy of rotation, as a sum of the two.

Thus, the calculation of the planet's rotational energy, which in balance equations plays a considerably greater role than the rotational moment of the planet, leads to the new conclusion that the thermal losses decrease with the increase of rotational velocity.

It is possible to find a planet rotation corresponding to maximum thermal losses. Let us assume that a certain set of particles and bodies is consolidating into a planet in two ways differing in their thermal losses. In the first case, the planet is formed on a circular orbit with the radius R. In the second case it is formed on an orbit with the radius R + δR. The full moment of the momentum should be identical; therefore, the change of the orbital moment is compensated by the change in the rotational moment /162

$$\delta K_r = - \delta K_0 = - \frac{m}{2} \sqrt{\frac{GM}{R}} \delta R,$$

This will lead to a change in the rotational energy

$$E_r = \frac{1}{2} I_r \omega_r^2 , K_r = I_r \omega_r; \delta E_r = I_r \omega_r \delta \omega_r = \omega_r \delta K_r;$$

$$\delta E_r = - \frac{m}{2} \omega_r \sqrt{\frac{GM}{R}} \delta R.$$

The change in the orbital energy

$$\delta E_0 = \frac{GMm}{2R^2}\,\delta R = -\frac{\omega_c}{\omega_r}\,\delta E_r$$

proves to be mush smaller than δE_r due to the smallness of the orbital angular velocity ω_c in comparison with the rotational velocity ω_r. Therefore, the change in the total mechanical energy

$$\delta E_0 + \delta E_r = -\frac{m}{2}\sqrt{\frac{GM}{R}}\,(\omega_r - \omega_c)\,\delta R \qquad (24)$$

is determined by the change of δE_r. If the thermal losses are increased, i.e., the mechanical energy is decreased, then at $\omega_r > \omega_c$, $\delta R > 0$ and the velocity of rotation will decrease.

The losses of energy are maximum when the sum of the orbital and rotational energies is minimum. For this, it is necessary that $\delta E_0 + \delta E_r = 0$, according to (24), that $\omega_r = \omega_c$. Hence, the thermal losses are maximum when the planet rotates about its own axes with the velocity of rotation of the cluster itself, i.e., when it does not rotate relative to the cluster. At the maximum thermal losses, the rotation would become direct but it would be too small in comparison with the actual rotation of the planets. The greater the rotation is relative to the cluster, the smaller the thermal losses. In addition, as far as the losses are concerned, it is quite immaterial in which direction, counting from ω_c, the rotation is occurring, as it is possible to show that the losses are identical for the velocity $\omega_c + \Delta\omega$ and for the velocity $\omega_c - \Delta\omega$. For $\Delta\omega > \omega_c$ the question of direction of rotation remains open.

The conclusion of direct rotation at sufficiently large heat losses was arrived at also by G. F. Khilmi (Ref. 16). It was based on an inequality, identically satisfied, and valid for any system of bodies independent of the character of their motion. /163

$$\frac{\Theta^2}{J^2} \leq 2T - \Sigma\, m_i r_i'^2 , \qquad (25)$$

where Θ and J^2 are respectively the orbital moment of the momentum and the moment of inertia of the system. The right side of this inequality is always greater than 0 since the radial velocities of the bodies r_i'

are smaller than their full velocities entering into the orbit kinetic energy of the system T. Therefore when $T \to 0$, the right side also approches 0. Hence, according to G. F. Hilmi, with an appropriate decrease in T, the orbital moment of momentum of the system should decrease, i.e., it should transform into a rotational moment. The thermal losses at collisions lead to the over-all decrease in T, and hence to direct rotation of the forming bodies.

However, the deduction of the substantial decrease in T and the direct rotation is justified only in the case of all the bodies in the system consolidating into one (central) body. But then, it is self evident, as the full moment of the system (not equal to 0) is transformed into the rotational moment. In the case of the formation of a system of bodies moving in circular orbits, a substantial decrease in T does not occur. The extinction of relative velocities of the bodies at collisions decreases the right side of the inequality to a magnitude $mv_z^2 \sim i^2T$. A change in T due to the consolidation of the bodies into a planet depends on the distance R_0 of the forming planet from the Sun. Prior to the consolidation of the bodies and particles into a planet, their orbital kinetic energy T is proportional to a_e^{-1}, and after the consolidation, $T \propto a_0^{-1}$. From (13) in which $\gamma < 0$ and on the order of $-\epsilon'$, it is seen that $a_0 \approx a_e$, i.e., that T changes only slightly. The condition for direct rotation (2) assumed by O. J. Schmidt $R_0 < R_e$ even leads to $T_0 > T$, i.e., not to a decrease in kinetic energy T at the consolidation of bodies into a planet, but, on the contrary, to its increase.

The planet will rotate the faster, the smaller its orbital moment, i.e., the smaller its orbital radius, and hence the greater the orbital kinetic energy T. Thus, the relationship between the orbital energy T and the rotation of the planet is actually reversed; this is in no way reflected in the inequality used by G. F. Hilmi. The very minute effect of the origination of rotation, connected with the transition of only one millionth part of the orbital moment into a rotational moment, can not be revealed in an inequality in which the difference between the right and left sides may reach more than 10%.

It must further be pointed out that the rotation energy of the planet does not appear as such in (25); it is combined with the energy of the losses. Because of that, an increase in rotational velocity would appear to be coupled with a decrease in thermal losses, as was deduced above by means of (24). It has the appearance of rotational energy /164 accounting for an economy in thermal losses. Thus, the relationship between rotation and thermal losses really also emerges as reversed.

In his subsequent book G. F. Hilmi (Ref. 17) analyzing the same inequality (25) imposes on the system an additional restriction. He assumes that the moment of inertia of the system J^2 decreases with time. However, even with this supposition it does not follow from the inequality (25) that direct rotation of the planets is a requisite of the process of their formation.

It is necessary to emphasize that all the objections cited above do not concern the bases of O. J. Schmidt's theory of formation of planets by means of accumulation of rigid bodies and the important role played in that process by the thermal losses. The objections are directed only against the idea that great thermal losses expressly conditioned the direct rotation of planets. The rotation of the planets as well as the thermal losses themselves were conditioned by specific concrete collisions of the consolidating bodies. The first is not a consequence of the second, since each is determine by basic rules of motion of the consoldating bodies. If, for example, the planets were forming in a non-rotating cluster they would not have acquired a direct rotation even if the losses reached magnitudes of the same order.

It is natural to connect the direct rotation of the planets with the general rotation of the entire cluster whose members are consolidating into planets. The bodies landing on the embryo of the planet actually may happen to impinge some to the right and some to the left of the axis of rotation bringing either a positive or a negative moment. However, the over-all rotation of the entire cluster of the bodies creates a certain small asymmetry of impacts which in the final analysis determines the rotation of the planet (Ref. 18).

The moment which is imparted to the planet's nucleus m by an impinging body of mass dm which prior to the encounter had velocity v relative to the planet's nucleus is equal to $\alpha r v dm$, where αr is the distance in line of aim of the impinging body (in a tangential impact $\alpha = \sqrt{1 + \frac{2\,Gm}{v^2 r}}$). The average moment which is imparted by many bodies will be determined by the sum of these vectors. Let $\bar{\alpha}$ be the average value of α unencumbered by random fluctuations arising at impingements by larger bodies. Then the planet's rotational moment increment coinciding in direction with the moment of the entire cluster is equal to

$$dK_r = \bar{\alpha} r v dm. \qquad (26)$$

The magnitude $\bar{\alpha}$ exactly characterizes the degree of asymmetry of the impacts. Arising from the fact of the planet's embryo and of the impinging bodies' motion around a central body - the Sun.

Assuming, as a first approximation, that $\bar{\alpha}$ remains constant /165 during the process of the planet's growth and taking for v the generally used value $v = \sqrt{\frac{Gm}{\theta r}} \propto r$, where r is the radius of the growing planet, then

$$dK_r \propto r^2 \text{ dim} \propto m^{2/3} \text{ dm and } K_r \propto m^{5/3}.$$

Hence, in the course of the entire process of growth

$$\omega \approx \text{const.} \tag{27}$$

It may also be assumed that the degree of asymmetry of the impingements determined by the rotation of the total cluster was essentially not dependent upon the distance of the growing planet from the Sun. This would lead us to the deduction of approximate equality of the periods of rotation for all planets, independent of their mass and distance from the Sun. It is known that one of the features characteristic of the solar system is precisely the small difference in periods of the planet's rotation despite the tremendous differences in their masses. Therefore, it follows that the above assumptions are close to actuality. The impingements by larger bodies caused deviations of the actual moments of the planets from the average, resulting in inclinations of equatorial planes of the planets to the central plane of the cluster. In view of the large magnitude of the moment that may be imparted by individual bodies, it must be concluded from the fact of the planet's direct rotation, that the masses of the largest impinging bodies during the final stage of the planet's growth were less than 10^{-2} of the mass of the planet itself. In the case of Uranus, this condition evidently was not fulfilled and the random component of the rotational moment turned out to be greater than the average.

Thus, it is possible to assume that the concept of the asymmetry of the impingments being the cause of the planet's rotation is basically correct. It leads to the deductions about the small variation in the periods of rotation of the planets and about the identical direction of their rotation, which it is more natural to assume as coinciding with the direction of rotation of the entire cluster than as counter to it.

This result is derived from the natural assumptions about the variation of the rotational moment in the nucleus-planet during the process of its growth. In a similar way it is possible to make certain most seemingly natural assumptions about the rotational energy of the planet's nucleus and then compare the obtained result with the previous

result. From the relationships (23) it is seen that the only simple assumption regarding E_r' which directly derives from dimensional considerations, is to allow that $T_r' \propto \frac{Gm}{r}$ since on the right side U_p and /166 v^2 are proportional to GM/r and since consequently, $E_t \propto \frac{Gm}{r}$. Then we obtain $dE_r \propto m^{2/3}\, dm$, i.e., $E_r \propto K_r\, \omega \propto m^{5/3}$, and since $K_r \propto m^{5/3}\, \omega$, we again come to the conclusion that $\omega \approx$ const.

The agreement of results arrived at by two different methods and also the general agreement with the actual data of the rotation of planets attests to their probability. However, it is necessary to emphasize that this result must still be regarded as a first approximation only. For a more rigorous solution of the problem concerning the planet's rotation, it is necessary to produce a quantitative evaluation of the magnitude of asymmetry $\overline{\alpha}$ based on a statistical examination of the restricted problem of three bodies.

However, this still would not have given a complete solution of all the problems connected with the planet's rotation. The following questions would still have remained unanswered.

1. Why are the orbits of the satellites located, as a rule, in the plane of the planet's equator? Is this related to the landing on the planet of a considerable quantity of matter from the satellite cluster (or is this connected with some other manifestation of the consolidation process), or is this related to the long dynamic evolution of the satellite systems that took place after their formation? If such a dropping of matter from the satellite cluster did occur, then in what way did it affect the rotation of the planets?

2. What was the initial rotation of the Earth like? There are serious arguments in favor of the idea that the Moon was at first at a considerably closer distance to the Earth than now, arguments deriving from considerations of tidal evolution, as well as from contemporary concepts of the process of the Moon's formation (Ref. 19). Did the recession of the Moon from the Earth occur because of tides or because of the Moon's consolidation with matter having a much greater moment of momentum relative to the Earth? In the first case, the Earth should have been initially rotating several times faster, but then it would have been rotating considerably faster than the other planets and would have thus, for some reason, deviated from the general rule.

3. Can the moment of the momentum of the Earth-Moon system be considered constant? E. Holmberg (Ref. 20), in particular, suggested the considerations that the semi-diurnal pulsations of the Earth's atmosphere could support the rotation of the Earth with a period close

to the period of the natural oscillations of the atmosphere. Observations reveal that the maximum compression (the protuberance) lags hours behind the Sun orientation. The tidal effect of the Sun on the atmospheric protuberances could, according to Holmberg, impart the necessary moment of momentum to the Earth. However, presently, the acceleration of the Earth's rotation by solar atmospheric tides is of an order less than its deceleration by lunar sea tides (Ref. 21) and the supposition that in the past, the acceleration was much greater is not very convincing. So far, there is still no meeting of minds on this question, and further investigations are necessary. /167

REFERENCES

1. H. Poincare. Lecon sur les Hypotheses Cosmogoniques (Lesson on Cosmogonic Hypotheses). Paris, 1911.

2. V. S. Safronov. Voprosy Kosmogonii (Problems of Cosmogony) 7, 121, 1960.

3. V. G. Fesenkov. Astron. Zh. 28, 492, 1951.

4. L. E. Gurevich and A. I. Lebedinsky. Izvestiya A. N. SSSR, Seriya Fizicheskaya 14, 765, 1950.

5. V. S. Safronov. Doklady A. N. SSSR, 130, 53, 1960.

6. V. S. Safronov and Ye. L. Ruskol. Voprosy Kosmogonii (Problems of Cosmogony), 5, 22, 1957.

7. K. E. Edgeworth. Monthly Notices of the Royal Astronomical Society, 106, 470, 1946.

8. Ye. L. Ruskol. Voprosy Kosmogonii (Problems of Cosmogony), 7, 8, 1960.

9. G. P. Kuiper. Extr. Mem. Soc. Roy. Liege 4, Ser., 13, 361, 1953.

10. I. S. Shklovskiy. Astron. Ah., 28, 234, 1951.

11. F. Hoyle. Monthly Notices Roy. Astron. Soc., 106, 406, 1946.

12. R. A. Lyttleton. Monthly Notices Roy. Astron. Soc., 121, 551, 1960.

13. F. Hoyle. Voprosy Kosmogonii (Problems of Cosmogony), 7, 15, 1960.

REFERENCES (Cont'd)

14. O. J. Schmidt. Chetyre Lektsii o Teorii Proiskhozhdenia Zemli (Four Lectures on the theory of the Origin of the Earth) 3rd Revised ed. Izd-vo A. N. SSSR, 1957.

15. V. S. Safronov. Izvestiya A. N. SSSR, Ser. Geofiz., No. 1, 139, 1959.

16. G. F. Hilmi. Problema n Tel v Nebesnoy Mekhanike i Kosmogonii (The Problem of n Bodies in Celestial Mechanics and Cosmogony). Doklady A. N. SSSR, 77, 589, 1951.

17. G. F. Hilmi. Kachestvennye Metody v Problemye n tel (Qualitative Methods in the Problem of n Bodies) Izd-vo A. N. SSSR, 1958.

18. V. S. Safronov. Voprosy Kosmogonii (Problems of Cosmogony) 7, 59, 1960.

19. Ye. L. Ruskol. Astron. Zh. 37, 690, 1960.

20. E. R. Holmberg. Monthly Notices Roy. Astron. Soc., Geophysical Supplement, 6, 325, 1952.

21. W. H. Munk and G. J. F. MacDonald. The Rotation of the Earth, Cambridge, Mass., 1960, p. 221.

ON THE DISPERSION OF VELOCITIES IN ROTATING SYSTEMS GRAVITATING BODIES WITH INELASTIC COLLISIONS /168

By

V. S.Safronov

The energy increase of irregular motion in a system of gravitating bodies of identical masses m and radii r with a differential rotation is evaluated by a well-known method of determining the heating of viscous medium in the presence of a velocity gradient. The expression for quantity of liberated energy which is applied in rotating systems having a great mean free path is derived. The inelastic collisions of bodies decrease the energy of irregular motion more than the relative velocities. From the conditions of the equality of the emitted and absorbed energy of irregular motion, the expression was derived for the (equilibrial) relative velocity of bodies in the system. This velocity has practically the same character of dependence on m and r as the equation found earlier by L. E. Gurevich and A. I. Lebedinskiy ($v^2 \propto Gm/r$). The coefficient of proportionality is determined by the degree of inelasticity of collisions, and it is also dependent upon the parameter which is connected with the characteristics of the energy and momentum transfer in rotating systems with a long duration of free flight. The numerical value of this parameter is not accurately known.

The problem concerning the dispersion of the velocities of gravitating bodies which are subjected to inelastic collisions presents a spontaneous interest in planet cosmogony since the process of accumu- /169 lation of planets primarily is reduced to collisions and junctions of a great number of protoplanet bodies. Having been formed in a planar dust disc, these bodies had initially small relative velocities and moved practically in Keplerian circular orbits. But in the course of their combination and increase in mass, they began to interact gravitationally with one another, and their relative velocities gradually grew, i.e., the eccentricities of their orbits grew.

An approximate expression for the dispersion of the velocities in protoplanets was formulated in 1950 by L. E. Gurevich and A. I. Lebedinskiy (Ref. 1). The period of the bodies' conjuncture is much less than their periods of revolution around the Sun. Therefore, the conjuncture might be considered as a two-body problem: the vector of the relative velocity v of the approaching bodies having masses m does not change in magnitude

and merely rotates through an angle of $\theta \approx \frac{Gm}{Dv^2}$ or $\theta \ll \frac{\pi}{2}$. Therefore, a change in the eccentricity e of the body's orbit occurs. This can be approximated by the equation

$$e\Delta e \approx \frac{Gm}{DV_c^2}, \tag{1}$$

where V_c is the circular velocity, and D is the target distance. According to the authors, if you further assume that there is a change in the orbit's eccentricity Δe at the closest conjunctures when D = 2r (r is the radius of the body) will by its order of magnitude equal the value of the eccentricity e, then from (1) it is possible to find e and consequently the relative velocity of the bodies

$$v \approx eV_c \approx \sqrt{\frac{Gm}{2r}}. \tag{2}$$

The conclusion is qualitatively correct. However, it is not sufficiently rigorous. The correlation (1) is only valid with small values of θ, i.e., during (distant) conjunctures. Since $\frac{\Delta e}{e} \approx \frac{|\Delta v|}{v} \approx \theta$, then in this case, $\Delta e \ll e$. But the authors applied (1) to the close conjunctures taking D = 2r and assuming that with this, $\Delta e \approx e$. The error that is introduced is not clear since with large values of θ, the correlations prove to be more complex and it is impossible to obtain from them an expression for v which would be analogous to (2).

The dispersion of the velocities in gravitating bodies in a system with a differential rotation occurs as a result of the transition of energy of regular motion into energy of irregular motion. The potential energy of the system with respect to the central mass decreases; the system is somewhat compressed in the direction of the axis of rotation. In the case of absolute elastic collisions their velocities would /170
increase all the while and would be no correlation of type (2). In an actual system with inelastic collisions, the dispersion of the velocities is determined by the balance between energies that are acquired during the conjunctures and energies lost during collisions. In the work (Ref. 1) the assumption $\Delta e \approx e$ when D = 2r is essentially an obscure expression of such a balance. The essential "characteristic" dimension should be actually on the order of 2r. However, equation (2) does not show the way in which the dispersion of the velocities is dependent on the character of the bodies' collisions from the degree of their inelasticity. Therefore, the problem must be approached in greater detail in order to find this relationship more clearly.

1. The Dispersion of Velocities During Short Periods of Free Flight

We will consider that all bodies have identical masses m and radii r. So far periods of free flight of bodies are short in comparison with the distance from the Sun (i.e., so far the bodies themselves are small) the growth of the dispersion of their velocities can be evaluated by ordinary formulae of hydrodynamics. In a flux possessing axial symmetry and rotating with an angular velocity of $\omega(R)$, the amount of energy dissipated in 1 cm^3/sec due to the molecular viscosity is determined by the equation

$$E = \eta R^2 \left(\frac{d\omega}{dR} \right)^2, \tag{3}$$

where $\eta \approx 1/3\ \rho v \bar{\lambda}$, is the coefficient of viscosity; and R is the distance from the axis of rotation. Having applied (3) to our system of bodies, we obtain the amount of energy which goes into the increase of their relative velocities. The correlation (3) has a simple physical interpretation. On the average, 1/3 of all particles travel in a radial direction. During the period of free flight path τ following the path λ they acquire a relative velocity of the differential motion $\Delta v = R \frac{d\omega}{dR} \lambda$ which transfers from a regular into a chaotic velocity. During the period $\tau = \lambda/v$ the thermal energy $1/2\ \Delta v^2 = \frac{1}{2} R^2 \left(\frac{d\omega}{dR}\right)^2 \bar{\lambda}^2$ is emitted per one unit of mass. Since $\bar{\lambda}^2 \doteq 2\bar{\lambda}^2$, then by having divided this expression by τ and multiplying it by $1/3\ \rho$, where ρ is the density of the medium we obtain (3).

So far only the bodies of small mass are being investigated, their gravitational interaction with one another is small and the period of free flight is determined by their geometric cross section. Between the two consecutive collisions the body acquires from the differential rotation an energy of the random motion which equals on an average one unit mass /171

$$\epsilon_1 = \frac{E}{\rho} \tau = \frac{1}{3} R^2 \lambda^2 \left(\frac{d\omega}{dR} \right)^2 . \tag{4}$$

Let us assume that during the collision, the body loses an energy of ϵ_2 per unit of mass. This energy amounts to a portion ζ of its

energy of relative motion. After having designated the velocity of the body after the collision by v_1 and the velocity before the following collision by v_2 we have

$$v_1^2 + 2\epsilon_1 = v_2^2 \; , \; 2\epsilon_2 = \zeta v_2^2 . \tag{5}$$

Then

$$\epsilon_2 = \frac{\zeta}{2} (v_1^2 + 2\epsilon_1) = \frac{\zeta}{2} (v^2 + \epsilon_1) , \tag{6}$$

where v^2 is the average value between v_1^2 and v_2^2. As a result of the total action of both effects, the body acquires per unit of mass per second the energy

$$\epsilon = \frac{\epsilon_1 - \epsilon_2}{\tau} = \frac{\zeta v^2}{2\tau} \left[\frac{2 - \zeta}{3\zeta} \tau^2 R^2 \left(\frac{d\omega}{dR} \right)^2 - 1\right] . \tag{7}$$

In Keplerian circular motion

$$R^2 \left(\frac{d\omega}{dR} \right)^2 = \frac{9}{4} \frac{GM_\odot}{R^3} = \frac{9\pi^2}{P^2} , \tag{8}$$

where P is the period of revolution around the Sun. In a system of identical bodies with a radius r, the geometric cross section of the collision is equal to $4\pi r^2$. But the collisions close to tangential have little effect. Therefore, we designate the actual cross section by $\xi\pi R^2$, where ξ is on the order of one unit. Then the time of the period of free flight between collisions is equal to

$$\tau = \frac{\lambda}{v} = \frac{4r\delta}{3\sqrt{2}\xi\rho v} , \tag{9}$$

where δ is the density of the body and ρ is the density in the cluster, i.e., the mean mass of the substance entering into the bodies per unit volume.

In planar rotating systems with the central field of gravity, the values of ρ and v are related with the surface density of the cluster and the period of rotation by the simple correlation (Ref. 2)

$$\rho v = \frac{4\sigma}{P} . \tag{10}$$

Therefore

$$\tau = \frac{P\delta r}{3\sqrt{2}\xi\sigma} . \tag{9'}$$

By substituting in (7) the expressions (8) and (9'), we find /172

$$\epsilon = \frac{\zeta v^2}{2\tau} \left[\frac{(2 - \zeta)\ \pi^2\delta^2 r^2}{6\zeta\xi^2\sigma^2} - 1\right] . \tag{11}$$

The analysis of this correlation allows the following conclusions to be made.

1. The velocities of the bodies decrease with time if the dimensions of the bodies are sufficiently small:

$$r < r_1 = \sqrt{\frac{6\zeta}{2 - \zeta}}\ \frac{\xi\sigma}{\pi\delta} . \tag{12}$$

If $r > r_1$ (all the bodies are assumed identical) then the velocities of the bodies increase regardless of v, i.e., they do not tend to approach a limit of the type (2). For the zone of the Earth, $\sigma \approx 10$ gm/cm^2; hence, r_1 is on the order of several centimeters.

2. The increase of velocities in the case being investigated occurs not because of the great gravitational interaction, i.e., "rocking" each other, but rather because of the system's differential rotation.

3. If, in the protoplanetary any dust layer, by some means a significant number of bodies with a radius greater than r_1 (comprising a large part of the layer's masses) succeeded in being formed, then the gravitational instability in the layer will not be able to occur. Actually, the increase of velocities when $r > r_1$ denotes a decrease in density, and if the density were less than the critical by $r \approx r_1$ then by its future decrease, it would be even less capable of achieving the critical value.

From this, it still does not follow that there was no gravitational instability in the dust layer. From (11), it is obvious when r is several times less than r_1 the first term in parentheses can be disregarded. Substituting τ from (9') we obtain the change in v in time

$$\frac{dv}{dt} = - \frac{3\xi\zeta\sigma}{\sqrt{2\delta P r}} v. \tag{13}$$

The velocity of the particle decreases by e times during the period $\frac{r}{r_1} P$, i.e., less than during the period of revolution around the Sun. Only the constantly acting outer excitations can rescue a cloud of small particles from the rapid descent. We have already mentioned earlier that such excitations can be expected in the inner part of the cloud where conditions for the occurrence of gravitational instability are more rigid.

Let us now evaluate the velocities of the bodies where $r > r_1$. Now in (11) it is possible to neglect the second term in brackets. Then,

$$\epsilon = v \frac{dv}{dt} \approx \frac{\pi^2 (2 - \zeta) \delta r}{2\sqrt{2}\xi P \sigma} v^2. \tag{14}$$

The velocities increase by e times during the period $\frac{r_1}{r} P$, /173 i.e., also during a period which is less than the revolution around the Sun. Let us assume that the bodies unite during collisions. Then one can find v relative to the dimensions of the bodies since

$$\frac{dm}{dt} = \sqrt{2}\gamma\xi\pi r^2 \rho v, \tag{15}$$

then on the basis of (10)

$$\frac{dr}{dt} = \frac{\sqrt{2}\xi\gamma\sigma}{\delta P}, \tag{16}$$

where γ is the probability that the bodies will unite during collisions, i.e., the magnitude on the order of unity. From (14) and (16) we obtain

$$\frac{dv}{v} = \frac{\pi^2 (2 - \zeta) \delta^2}{4\gamma\xi^2\sigma^2} r\, dr. \tag{17}$$

Let us assume that when $r = r_0 > r_1$ velocity of the bodies $v = v_0$. Then

$$\ln \frac{v}{v_0} = \frac{\pi^2 (2 - \zeta)\delta^2 r_0^2}{8\gamma\xi^2\sigma^2} \left(\frac{r^2}{r_0^2} - 1 \right). \tag{18}$$

When $r_0 = 5$ cm, the multiplier before the bracket is on the order of unity. Therefore, from (18) it follows that it is sufficient that r would increase only several times, such that v would reach very high values and would lead to the destruction of the bodies as such. However, such deduction would be invalid. With an increase in v, the thickness of the layer increases, and hence the length of the period of free flight of the bodies increases. But in the presence of large values of λ which comprise the basis of this deduction the correlation (4) ceases to be applicable. During the intervals between collisions, the bodies under influence by solar gravity travel along elliptical orbits. Therefore, with any suitably large values of λ the maximum displacement of the body by R does not exceed 2ae, i.e., about 2eR if the eccentricities e of the orbits are not great. When $\lambda > eR$, i.e., when $\tau > \frac{1}{4}$ P, correlations (4) and (7) reveal an increased value for ϵ. This leads to an excessively rapid growth of v in (18). From (9') and (12) we have

$$\tau = \frac{1}{3\pi} \sqrt{\frac{3\zeta}{2 - \zeta}} \frac{r}{r_1} P. \tag{19}$$

From this it is evident that τ reaches a magnitude of 1/4 P when the value r is only slightly greater than r_1. Hence, correlation (18) ceases to be correct very early and has to be substituted by another value suitable for large values of λ.

2. The Dispersion of Velocities During Long Periods of Free Flight /174

In rotating systems during the transition to longer periods of free flight paths the character of the substance and motion transfer essentially changes. Thus, in the semi-empirical theory of Prandtl for turbulent motion which is based on the concept concerning the average course of intermixing, the tangential directions are determined not by the gradient of the angular velocity but by the gradient of the momentum of vorticity [see for example von Karman (Ref. 3)]. Then correspondingly, also the expression for the thermal energy being emitted should change. Actually,

in the intervals between collisions, the body moves with a constant momentum of vorticity. Let us utilize the same principle on the basis of which expression (3) was obtained. The body possessing at a distance R from the Sun an average motion in the direction of rotation ($v_\theta = 0$) after being shifted into a point at a distance $R + \Delta R$ will have systematic velocity in the direction of rotation relative to the circular velocity

$$v_\theta = - \frac{1}{R} \frac{d\ (\omega R^2)}{dR} \Delta R. \tag{20}$$

The maximum possible value ΔR equals eR and the average square $\overline{\Delta R^2} = \frac{1}{3} e^2 R^2$. Therefore

$$\frac{1}{2} \overline{v_\theta^2} = \frac{e^2}{6} \frac{d\ (\omega R^2)}{dR} . \tag{21}$$

Let τ_s be the time between the two consecutive collisions of bodies and τ_g the time of the bodies' energy exchange caused by their conjuncture, i.e., the period of relaxation. Then the period of free flight is determined from the condition

$$\frac{1}{\tau} = \frac{1}{\tau_s} + \frac{1}{\tau_g} . \tag{22}$$

By applying the same arguments from which expression (3) was derived one can consider that within the time τ the energy of the random motion per unit of mass increases by a magnitude of $1/2\ \overline{v_\theta^2}$ which is determined according to (21). However, Prandtl's theory which leads to the gradient of the momentum of vorticity is not entirely rigorous. According to Vasyuginskiy (Ref. 4) the tangential stresses are determined by a more complex expression in one extreme case which, when the exchange is isotropic, the stresses are determined by the gradient of angular velocity and in another extreme case when the exchange is entirely radial. The exchange is determined by the gradient of the momentum of vorticity (according to Prandtl). Actually, the purely radial exchanges do not materialize and a certain intermediate case should take place. /175
In Keplerian rotation $\omega \propto R^{-3/2}$, $\omega R^2 \propto R^{1/2}$. The tangential stresses calculated by the gradient ωR^2 are three times less than according to the gradient ω. With respect to this, the correlation (21) indicates that the energy is nine times less than the energy which we might have obtained if, similar to (3) we determined it from the square of the gradient of angular

velocity. From the previous discussion it follows that in (21) it is probably necessary to introduce a certain constant correction factor. The determination of this factor goes beyond the scope of the present article. We will confine ourselves to the supposition that within time τ the average energy of the random motion of a mass unit is increased by the value of $1/2\ \beta\overline{v_\theta^2}$ where β is on the order of unity. The energy acquired during the time τ_s will be equal to*

$$\epsilon_1 = \frac{\beta}{2}\overline{v_\theta^2}\left(1 + \frac{\tau_s}{\tau_g}\right). \tag{23}$$

As before, we will consider that, during a collision, energy ϵ_2 determined by expression (6) is lost,

$$\epsilon_2 = \frac{\zeta}{2}\left(v^2 + \epsilon_1\right).$$

The complete change of energy of the random motion in 1 sec will be equal to

$$\epsilon = \frac{\epsilon_1 - \epsilon_2}{\tau_s} = \frac{1}{\tau_s}\left[\frac{(2 - \zeta)\ \beta e^2}{12}\left(\frac{d\ (\omega R^2)}{dR}\right)^2\left(1 + \frac{\tau_s}{\tau_g}\right) - 1\right]. \tag{24}$$

---

*Edgeworth (Ref. 5) with the help of a correlation approximating (3) attempted to evaluate the increase of dispersion of stellar velocities. Having correctly noted that in the case of large periods of free flight the magnitude λ in the formula should be substituted by the amplitude of oscillations of the star along the radius vector ΔR, he still admitted an error by asserting that it is sufficient to produce a substitution in the coefficient of viscosity $\eta \propto v\lambda$. Actually, it is seen from the conclusion (3) and also from (4) and (7) that the exchange should be made not in the expression $v\lambda$ but rather in the expression λ^2/τ. Hence, instead of $v\Delta R$, it should be $\Delta R^2/\tau = v\Delta R \cdot (\Delta R/\lambda)$. λ is five times larger than ΔR for the approach of separate stars. Although the period of increase in the velocity dispersion due to stellar conjunctives ($\sim 10^{10}$ years) which was found so many times by Edgeworth has proved to be less than what is actually necessary. Only conjunctures of stars with significantly more massive objects ($M \sim 10^6 M_\odot$) can lead to the cosmogonically acceptable rate of growth in the dispersion of stellar velocities.

We will assume τ_g equals period of relaxation according to Chandrasekar (Ref. 6)

$$\tau_g = \overline{T}_E = \frac{1}{16}\sqrt{\frac{3}{\pi}}\,\frac{(\overline{v^2})^{3/2}}{G^2 m\rho\ \ln\left[1 + \frac{D_0 v'^2}{2Gm}\right]} . \tag{25}$$

Here, D_0 is the average distance between the bodies and $v'^2 \approx 2v^2$ where /176
v' is the relative velocity of the bodies, v is the velocity of the body relative to the centroid. The time between collisions not counting the mutual attraction of the bodies is determined according to (9). The gravitation of the bodies decreases τ by $(1 + \frac{Gm}{v_2^2 r})$ times where v_2 is the velocity of the body relative to the centroid before the collision. From (5) and (6) we have

$$v_2^2 = v^2\,\frac{2}{2 - \zeta} .$$

Consequently,

$$\tau_s = \frac{4r\delta}{3\sqrt{2}\xi\rho v\left[1 + \frac{Gm\,(1 - \zeta/2)}{v^2 r}\right]} , \tag{26}$$

$$\frac{\tau_s}{\tau_g} = \frac{16\,\pi G^2\delta m r\ \ln\left(1 + \frac{D_0 v^2}{Gm}\right)}{3\xi(\overline{v^2})^2\left[1 + \frac{Gm\,(1 - \zeta/2}{v^2 r}\right]} . \tag{27}$$

Now we shall find the value v in the presence of which ϵ is reduced to 0. Since

$$\frac{d\,(\omega R^2)}{dR} = \frac{1}{2}\,v_c \text{ and } e^2 V_c^2 \approx \frac{4}{3}\,v^2, \tag{28}$$

then the right term (24) is reduced to 0 if

$$\frac{\tau_s}{\tau_g} = \frac{18\zeta}{(2 - \zeta)\beta} - 1. \tag{29}$$

The comparison (27) and (29) leads us to the conclusion that the dependence of $\overline{v^2}$ on the masses and the radii of the bodies has the form

$$\overline{v^2} = \frac{Gm}{\theta r}, \tag{30}$$

where

$$\theta^2 = \frac{18\,\xi\zeta - \xi\beta\,(2 - \zeta)}{4\,(2 - \zeta)\,\beta\,\ln\left(1 + \frac{D_0}{\theta r}\right)}\left[1 + \left(1 - \frac{\zeta}{2}\right)\theta\right]. \tag{31}$$

Thus, the dispersion of the velocities in protoplanetary bodies is determined by their masses and radii and also by the degree of the inelasticity of their collisions.*

In addition, the average distance between the bodies D_0 enters /177 into the expression for θ; however, the dependence on D_0 is extremely slight (the root of the logarithm). Therefore, θ can be considered practically constant in the growth process of increase if a perceptible change in ζ doe not occur.

Let us evaluate the magnitude ζ for the case where the colliding bodies are united i.e., for the positively inelastic collisions. Let us assume that two bodies of identical mass m have identical velocities v under the angle ψ of their relative directions. After the collision and junction of bodies, amount of motion should remain

---

*If, instead of T_E the time between close approaches which is $16\ln\left[1 + \frac{D_0 v^2}{Gm}\right]$ times greater, it is taken then θ^2 increases in the same way and θ is increased approximately ten times. This value of θ should be considered as obviously overstated since in the calculation distant conjunctures are not taken into consideration.

$$2\ mv' = 2mv \cos \frac{\psi}{2}\ .$$

this means that, after the collision, the velocity is equal to

$$v' = v \cos \frac{\psi}{2}\ . \tag{32}$$

According to the definition of ζ

$$\zeta v^2 = v^2 - v'^2\ .$$

Hence,

$$\zeta = \sin^2 \frac{\psi}{2}\ . \tag{33}$$

If vectors v are randomly distributed and the probability of collisions of bodies is proportional to their relative velocity Δv

$$\Delta v = 2v \sin \frac{\psi}{2}\ , \tag{34}$$

then the average value of $\bar{\zeta}$ is determined from the expression

$$\bar{\zeta} = \frac{\int \zeta \Delta v\ d\Omega}{\int \Delta v\ d\Omega} = \frac{\int_0^\pi \sin^3 \frac{\psi}{2} \cdot 2\pi \sin \psi\ d\psi}{\int_0^\pi \sin \frac{\psi}{2} \cdot 2\pi \sin \psi\ d\psi} = \frac{3}{5}\ . \tag{35}$$

Actually, the distribution of velocities of bodies is not isotropic.

In a radial direction the velocities on the average are twice as large as the tangential.

However, $\bar{\zeta}$ does not essentially change with this. Since the conclusion is attributed to absolute inelastic impacts the found /178 value of $\bar{\zeta}$ should be considered as the upper limit of ζ.

Let us now evaluate D_0. According to Chandrasekarv (Ref. 7)

$$D_0 \approx 0.55\ n^{-1/3}, \tag{36}$$

where n is a number of bodies per unit volume:

$$n = \frac{\rho}{m} = \frac{4\sigma}{Pvm} .$$

Substituting this value of n in (36) and substituting v according to (29) and after performing some simple calculations we obtain

$$\frac{D_0}{r} \approx 0.125 \sqrt[3]{\frac{P\delta r}{\sigma} \sqrt{\frac{\delta}{\theta}}} . \tag{37}$$

For the zone of the Earth $\sigma = 10$, $P = 3 \cdot 10^7$. With $\delta = 2$ and $r = 10^5$ we obtain $D_0/r \approx 10^3\theta^{-1/6}$. If we assume that for the zone of Jupiter $\sigma = 40$, $P = 3.7 \cdot 10^8$, $\delta = 2$ and $r = 10^7$, then we obtain $D_0/r \approx 10^4\theta^{-1/6}$. Thus, D_0/r proves to be on the order of 10^3 to 10^4. The magnitude under the logarithm in (31) is equal to $D_0/\theta r$. Substituting in (31) $\beta = 1$, $\xi = 1$ and instead of ζ the value which we got above $\bar{\zeta} = 3/5$, we find that $\theta = 0.56$ when $D_0/r = 10^3$ and $\theta = 0.47$ when $D_0/r = 10^4$. Thus it is to be expected that θ is on the order of 1/2. For a more accurate estimation of θ it is necessary to amke the coefficients β and average value $\bar{\zeta}$ more accurate.

The obtained result is attributed to the ideal case of the system of identical bodies which combine during collision. If the bodies during collisions do not combine their velocities will be greater. On the other hand the presence of dispersed substance (gas and small particles) in a cluster decreases the velocities of the bodies. In a system of unidentical bodies the picture will be still more complex. The larger bodies should have significantly less velocities since they deviate less under gravitational conjunctures with other bodies and they curve their own orbits to a great degree during the fall of smaller bodies on them. However, in order to examine the dependence of the velocities of the bodies on their masses it is necessary to examine the result of concrete conjunctures in detail.

The hydrodynamic method applied earlier is insufficient for this. In addition, it is necessary to know the function of the bodies' distribution in dimension. Essentially, these problems are closely related and should be considered simultaneously during the rigorous statement of the problem.

REFERENCES

1. L. E. Gurevich and A. I. Lebedinskiy. Izvestiya Akademii Nauk, SSSR Seriya Fizicheskaya 14, 765, 1950.

2. V. S. Safronov, Voprosy Kosmogonii (Problems of Cosmogony) Astron. Zh. 31, 499, 1958.

3. T. Karman. Symposium "Problemy Kosmicheskoy Aerodinamiki" (Problems of Cosmic Aerodynamics) Izdatel'stvo Inostrannoy Literatury, 1953, p. 157.

4. J. Wasiutynski. Studies in Hydrodynamics and Structure of Stars and Planets, Oslo, 1946.

5. K. Edgeworth. Monthly Notices of the Royal Astronomical Society, 108, 417, 1948.

6. S. Chandrasekarv. Printsipy Zvezdnoy Dinamiki (Principles of Stellar Dynamics), Izd-vo Inostrannoy Literatury, 1948, p. 72.

7. S. Chandrasekarv. Stokhasticheskie Problemy v Fizike i Astronomii (Stochastic Problems in Physics and Astronomy, Izd-vo Inostrannoy Literatury, 1947.

A BRIEF REVIEW OF RESEARCH ON MORPHOLOGICAL CHARACTERISTICS OF DIFFUSE NEBULAE /180

By

R. Ye. Gershberg

A brief review is given on the investigations of morphological characteristics of diffuse nebulae. The fundamental physical ideas connected with the interpretation of three very important types of structural details of diffuse nebulae are examined: bright rings on the border of luminous and dark nebulae; peripheral structures of the emission nebulae; and stretched and fibrousness formations of luminous and dark diffuse nebulae.

The relaxation period of all known elementary processes which determine the physical state of the inter-stellar matter, the recombination and ionization time, the period for establishing Maxwell distribution of the velocities of the particles with identical masses etc, are usually several times less than the period for the existence of cosmic objects in a fixed condition or the time of existence of a certain macroscopic motion. This essentially simplifies the investigation of the physical state in inter-stellar bodies since the stationary objects are investigated. Furthermore, since the radiation of the inter-stellar medium reflects its state only at a given moment, then the study of the evolution of the diffuse matter by its luminosity proves to be impossible. For the investigation of the evolution of inter-stellar medium it is necessary to turn to the properties which have a longer characteristic change period than the period of relaxation of the elementary processes. The consideration of the kinematic characteristics which are being determined directly from observations with known /181 physical properties of the diffuse matter renders it possible to construct dynamic theories, and then evolutional hypotheses on the development of various inter-stellar objects. The study of the structural characteristics of diffuse nebulae yields the criteria supporting the dynamic theories and is an important link in the construction of the over-all picture of the development of diffused matter in the galaxy and its relation to the stars. At first the problem of investigating the structure of diffuse nebulae for the clarification of their dynamics and evolution was accurately formulated by G. A. Shayn and V. F. Gaze (Ref. 1) in their first work dedicated to the study of the inter-stellar medium.

A systematic study of morphological characteristics of diffuse nebulae began comparatively recently. This is quite natural if you take

into consideration that the theory of physical processes in the inter-stellar medium in the first approximation was completed only in the beginning of the 1940's. At that time Stremgren (Ref. 2) showed that in the conditions of the inter-stellar medium the regions of ionized and neutral hydrogen should be rather abruptly defined in space. Up to this time the investigation of the structure of diffuse nebulae was reduced to a simple description of the odd forms or the classification of the nebulae on the basis of their external form. Such a classification did not always prove to be physically justified. Not stopping at the numerous descriptive works of this type [(a detailed bibliography can be found in the monographs of B. A. Vorontsov-Velyaminov (Ref. 3) and Cederblad (Ref. 4)] we shall examine only the fundamental physical ideas connected with the interpretation of three important types of structural details of diffuse nebulae: 1) the bright rings on the border of luminous and dark nebulae, 2) the peripheral structures of the emission nebulae, and 3) the stretched and fibrous formations in bright and dark diffuse nebulae.

1. The bright rings on the borders of luminous and dark nebulae are visible on certain photographs which were taken at the end of the 19th century (Refs. 5-7). Duncan (Ref. 8) while describing the photographs which were taken with the 100-in. reflector first turned his attention to these bright rings as being a characteristic detail of the diffuse nebulae. In regard to IC 434 he assumed that the bright rim around the "Horse Head" nebula (in Orion) was a result of the skimming of the bright substance by dark material.

In 1937 Struve (Ref. 9) composed the first list of the nebulae with bright rims and found that the brightest parts of the rings "point out" the stars which excite the nebulae. He also found that from the side of the dark nebula the brightness of the rings drops more steeply than from the side of the bright nebula and that all the luminous nebulae which are bordered by bright rings (in Struve's list there were ten such objects) have emission spectra. However, Struve considered that /182 the rings themselves belonged to the dark matter "the silvery edge of the clouds" which is projected on the other bright substance. In his opinion the radiation of the stars penetrates the dense clouds comparatively little because of the high coefficient of absorption of this radiation. From the front these illuminated layers are hardly visible because of their small surface brightness, but from the side they have the appearance of narrow sickles or bright rings, since the dispersed radiation is only slightly absorbed and the entire thickness of the bright substance is visible.

In 1946 Öort (Ref. 10) showed that during the collision of clouds of inter-stellar gas shock waves are formed which can cause a luminescence of matter whereby the illuminated zone should be quite narrow. He

proposed to clarify the bright rings by means of the collision of dense and rarefied clouds; the rarefied cloud is heated to a high temperature and begins to shine while a more dense cloud remains dark. Thackeray (Ref. 11) soon afterward completed special spectral photometric observations on the bright rings in three known diffused nebulae in the southern sky -- NGC 6523, NGC 6514, and NGC 6611. These observations revealed the emission spectrum for the bright rings; and in this way they showed that the model of "silvery edge of the dark cloud" was invalid. However, these observations were insufficient for the solution of the problem whether the bright rings are excited by a radiative or shock mechanism.

To the problem on the nature of bright rings is closely related the problem of the formation of dark "elephant trunks", i.e., the long and narrow introduction of dark matter into the bright nebulae, since these "trunks" are normally bordered by bright rings. By investigating the large direct photos Thackeray (Ref. 11) came to the conclusion that the elephant trunks are formed because of the motion of the dark substance in the direction of a star which excites luminescence of the emission nebulae. Rozhkovskiy (Ref. 12) came to the same conclusions. The hydrodynamic model of such a flux was given by Spitser (Ref. 13). Spitser proposed the examination of the dark "elephant trunks" as a result of the so-called Rayleigh-Taylor instability on the border of two media with a diverse density of matter and with an acceleration directed away from the denser medium to the less dense. Frieman (Ref. 14) and Layzer (Ref. 15) made calculations on such a motion (applicable to conditions of the inter-stellar medium). Not long before this van de Hulst (Ref. 16) proposed a different model of the formation of "elephant trunks" and bright rings: if an O-star can be born in a medium with certain fluctuation density, then the original region of the hydrogen ionized by the star has an irregular outline. The more dense regions are ionized at a lesser depth and in these places the border of ionized substance forms protuberances toward the star. The high pressure of the ionized gas compresses this protuberance and makes it more dense, whereby the /183 edge of the compressed area can be ionized and will be seen as a bright ring embracing the dark "trunk".

Soon after this Osterbrock (Ref. 17) investigated the "elephant trunks" in all the emission nebulae with distances which were reliably determined (such proved to be ten) and came to the following conclusions: the structures which are being observed in general conform to the Rayleigh-Taylor instability model although they have certain effects which were not seen in theory; the density of the bright rings essentially exceeds the mean density of substance in the emission nebulae which are bordered by rings. The evolution of the "trunks" can be concluded by a formation by an isolated "droplet", the dark globule. At first these small dark objects on the background of the emission nebulae were discovered by Bok

and Reilly (Ref. 18) and V. G. Fesenkov and D. A. Rozhkovskiy (Ref. 19) came to a conclusion concerning the physical relation of the globules with the dark matter which surrounds the emission nebulae.

At almost the same time as Osterbrock, Pottasch (Ref. 20) conducted a detailed statistical investigation of bright rings in 34 diffuse nebulae and discovered some new important laws. It was explained that the emission nebulae which contained bright rings were excited by stars of a more recent spectral class than O9. The position of the bright ring in the nebula is dependent on the spectral class of the exciting star i.e., rings are located closer to the hotter stars. The form of the bright rings changes with its distance from the exciting star from almost planar remote rings to firmly drawn rings and circular globules which do not have an apparent connection with the dark matter and are comparatively close to the center of the nebula. The planar rings are the largest and least dense, whereas the bright rings which border the narrow "elephant trunks" and the globules are the smallest and most dense. The decline in density of matter from the ring to the central parts of the nebula is maximum in the drawn-out nebulae with low density and at a minimum in the compact and dense nebulae. Analysis of these laws led Pottasch (Ref. 21) to the conclusion that the bright rings cannot be a result of the Rayleigh-Taylor instability, but rather they correspond to a picture of the extent of the ionization front in the inter-stellar medium with fluctuations in density. By using the theory of this process developed by Kan (Ref. 22) Pottasch showed that a configuration of bright matter occurs when a region of increased density encounters the ionization front. This configuration yields the exact course in brightness with distance from the exciting star as Osterbrock discovered in a typical bright ring. After this Kan (Ref. 23) theoretically examined the /184 problem concerning the stability of the ionization front and came to the conclusion that the "elephant trunks" are not a result of Rayleigh-Taylor instability since a similar instability should have led to the existence of lesser "excitations" than those brought about by the observed rings, and in addition the absorption in the ring itself yields a strong effect of the attenuation of instability.

Evidently, the model assumed by van de Hulst corresponds to the phenomenon of the bright rings. Recently Hershberg (Ref. 24) showed that the border of the ionized region will have significant protuberances and depressions with the fluctuations of density in the inter-stellar matter amounting to only 20 to 30% of the average density. These deformations are not caused by the drop in matter density immediately near the border of the zone H II, but by the fluctuations of density in the internal parts of the ionized region. In other words, the ionized front does not "encounter" any ready in the form of local matter condensations and does not bend by "flowing around" it, but was bent from the beginning. This

deformation of the ionization front and subsequently the hydrodynamic rupture leads to the formation of a local condensation. The design of future development of the bright rings assumed by Pottasch is valid in the given case.

The investigation of the electron temperature in the bright rings of diffuse nebulae NGC 7000 and NGC 6523 has shown (Ref. 25) that the temperature in these formations, in effect, do not differ from the electron temperature of more inner parts of the corresponding nebulae. Hence, the physical condition of the substance in these details cannot be perceptibly distinguished from the condition of substance in the usual H II regions and the bright rings can only be the result of increased density in the luminescent matter. These data obtained from observations contradict the concepts concerning bright rings as well as the information on the luminescent fronts (Oort), and also the concepts concerning the fact that these regions of the diffuse nebulae are excited by the more rigid ultraviolet radiation than the remaining parts of the ionized regions (Pottasch).

2. For a long time the large nebula in Unicorn NGC 2237 with the characteristic concentration of matter on the periphery proved to be the only object among the diffused nebulae. The unique nature of this nebula became noticeable when Minkovskiy (Ref. 26) determined its mass of $\sim$ 10,000 solar masses. However, in 1951 after having investigated the numerous photographs of the emission nebula obtained with illumination cameras in Cimeiz, Shayne and Gaze (Ref. 27) have shown that the peripheral structures, i.e., the solid external shells, the rings and individual arcs are sufficiently characteristic peculiarities of diffuse nebulae. Afterward, they found that another well-known /185 nebula NGC 6523 also has interesting peripheral details and is comparable by mass to NGC 2237 (Ref. 28). Somewhat later, they supplemented the initial list of peripheral diffuse nebulae; among them were objects from the extragalactic systems (Ref. 29). In 1953, Shayn and Gaze (Ref. 30) came to the conclusion that the peripheral structures cannot be equilibrial static formation but rather are dynamic configurations as a result of the influence of some kind of force on the diffuse matter which initially was found in the more central regions. Shayne and Gaze examined three possibilities for the formation of peripheral structures: "the formation of a shell as a result of the effect of radiation pressure as a single factor which determined the transformation of the nebula initially amorphis in structure into a peripheral nebula, the formation of a shell or ring as a result of the ejection of gas from one or several stars, i.e., super nova, and the formation of a shell or ring as a result of the motion of gas outward in a non-stationary formation of the motion stipulated by some reason during the period of the formation of the nebulae". Shayn and Gaze considered the last possibility the most probable. Subsequently the concepts concerning the transformation of

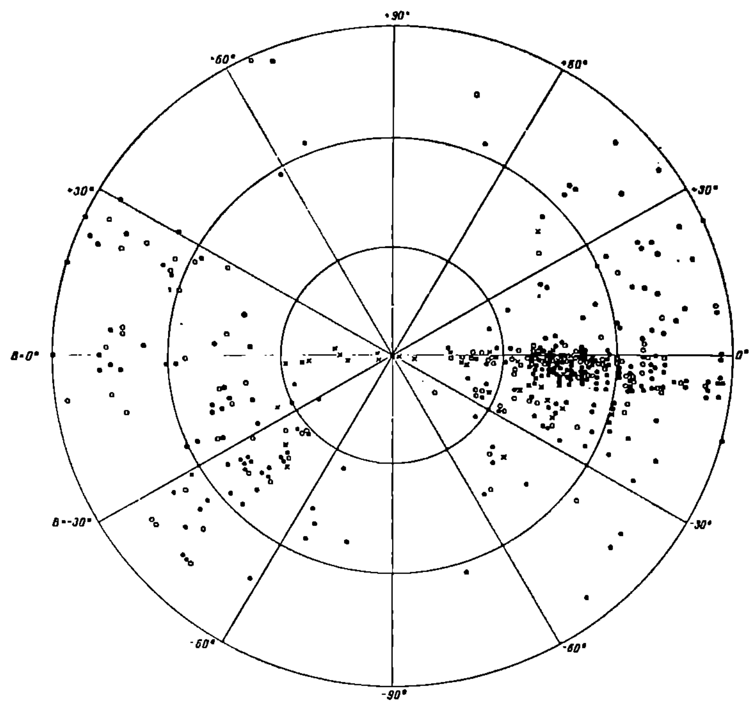


Fig. 1. Meridional Cross Section of a Sphere with a Radius of 15 Mparsecs. The black circles are spiral galaxies brighter than -17^m of absolute magnitude; the bright circles are elliptical galaxies; the crosses are spiral and irregular galaxies of a magnitude less than -17^m. The system of distances being used here can be characterized by the fact that the distance to the cluster in Virgo, after the consideration of light absorption is equal to 7.6 Mparsecs. The northern galactic hemisphere is located to the right. Scale of the illustration: 1 cm = 2 Mparsecs. On the cross section are projected all galaxies which deviate from L = 105° not more than by 15°.

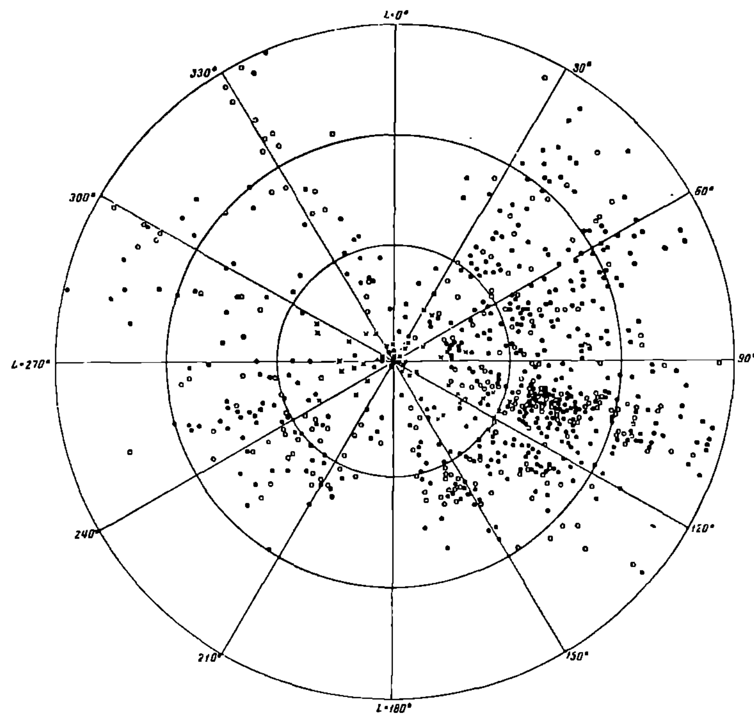


Fig. 2. Equatorial Cross Section of a Sphere with a Radius of 15 Mparsecs. On the cross section are projected all galaxies, which are located in the near equatorial layer with a thickness of 2 Mparsecs. The northern galactic hemisphere is shown to the right. The scale and designations are the same as on Fig. 1.

the amorphis into peripheral nebulae became an essential point in the over-all field of ideas concerning the evolution of diffuse nebulae, their expansion and deterioration (Refs. 31, 32). The numerous proofs presented by Shayn and Gaze favor the expansion and deterioration of the diffuse nebulae, i.e., the motion of gas in NGC 1976 relative to the Trapezium Orion, the existence of groups of the emission nebulae, the validity of distributions of exciting stars relative to the diffuse nebulae, etc., are so persuasive that evidently there is no need to doubt the existence of this process.

The expansions of the regions H II at first were considered theoretically by Öort and Spitzer (Refs. 33, 34). They considered the expansion of the emission nebulae as a mechanism which transmits parts of the radiation energy of hot stars to the clouds in the inter-stellar medium. Savedoff and Green (Ref. 35) calculated a hydrodynamic model with this type of motion. But those and others were only interested in the effects of expansion in neutral hydrogen and they did not consider the evolution of the emission nebula itself. Based on the Öort-Spitzer hypothesis Pronik (Ref. 36) proposed the following system for the formation of the peripheral structures. The expanding zone H II results in a shock wave in the surrounding gas (H I). The shock wave sets into motion a certain layer of cold compressed gas which circles the zone H II. During the expansion of zone H II this optical density declines and the compressed neutral hydrogen is ionized. If the density of the /186 inter-stellar gas as it gradually departs from the exciting star falls, then the ionization of the substance from the inner part of the layer of the compressed H I - gas can proceed faster than the "skimming" of new masses from the external area, so that in time the entire dense "barrier" will be ionized and will create a bright mantle around the remaining mass of the diffusing nebulae. This model explains well the observed structure of a diffusing nebulae NGC 6523 but being purely qualitative, it cannot substitute the theory of the formation of peripheral structures.

Recently calculations on the expansion of zone H II were conducted by this author (Ref. 37). From these calculations a new model for the formation of peripheral structures in diffusing nebulae follows. Expansion of the zone H II brings a shock character into the surrounding gas. This hydrodynamic motion is complicated by secondary effects: the expansion of the initial Stremgren zone leads to the decrease of its optical thickness in the Lyman continuum. Therefore, due to the shock wave along the compressed neutral gas an ionization skip is emitted. Since the shock wave which has passed through the neutral substance increases the pressure in the region H I to a magnitude which equals the pressure in the region H II, then on both sides of the secondary ionization skip the pressure is identical. From a hydrodynamic point of view the secondary ionization front evokes the phenomenon which corresponds to a surface explosion. The surface explosion emits distortion in both

directions. An approximate calculation of such a motion shows that the distortions which are emitted by the ionization front in the direction of the hot star can lead to a certain increase in density of matter in the outer parts of the zone H II. This increase in density is sufficiently great to explain the formation of peripheral structures.

The calculations (Ref. 37) revealed the inaccuracy of some assumptions in the Öort-Spitzer hypothesis. First of all the Stremgren zone is formed around the newly excited hot star and only thereafter the hydrodynamic interaction of hot (ionized) and cold (neutral) gases begins. Up to the formation of the Stremgren zone the ionization front was freely emitted. This did not result in hydrodynamic motions. But after the occurrence of this zone the further advancement of the ionization front is fully determined by the expansion of the region H II since the expansion of the initial Stremgren zone determines also the force of the ionization flux and the state of the substance which is affected by this ionizing radiation. Thus the hot star indirectly brings into motion the inter-stellar gas through the expansion of the ionized gas which surrounds it. This circumstance reduces the effectivity of the "rocket mechanism" of the acceleration of inter-stellar clouds to zero. This mechanism /187 was proposed by Öort and Spitzer. Furthermore, if the shock wave, which expands through a neutral substance, reaches regions with an abrupt drop in the density of the medium, then a rarefaction wave will affect the compressed gas (H I) and the frontal area of the compressed neutral gas will travel with the speed of gas dispersion in a vacuum. After the rarefaction wave encounters the ionization front the velocity of the ionization front increases relative to the compressed neutral gas, but the speed of dispersion of the gas H I in a vacuum will remain less. It is, therefore, highly improbable that the fluctuations of density of the neutral substances, which surround the initial Stremgren zone, might lead to the overtaking of the shock wave by the ionization front. This was assumed in (Ref. 36) for the explanation of observed peripheral structures.

3. We will now consider such an important structural peculiarity of the diffused nebulae, as for instance elongation and fibrousness. This peculiarity is encountered in objects which are extremely diverse in their nature as well as in their dimensions: the bright fibers in the Cancer nebula in NGC 6960--6992--5 and in other remnants of clouds of a super nova; objects of the "cosmic gas flux type", such as the nebula NGC 1499; the elongated striated gas fields in Cygnus, the system of fibres in the nebula NGC 2264; the elongated dark nebulae in Taurus and Serpents, the fine fibrous formations in the reflecting nebulae, for example in Pleiades. By no means is it evident that the formation of all of these structures is related to the effect of one and the same mechanism.

The firmly elongated nebulae are easily seen on the photos of stellar fields in the Barnard (Ref. 5), Ross and Calvert (Ref. 38) Atlases. These objects attracted attention before many elongated and fibrous emission nebulae were known. It was once considered, that the elongated dark nebulae were formed from more or less compact objects under the influence of tidal forces, as a result of the differential galactic rotation (Ref. 39). Shayn came up with a different hypothesis: The expansion of gaseous clouds (ionized as well as neutral hydrogen) due to the inner thermal and turbulent motions in a magnetic field can lead to the formation of such large elongated structures as the dark nebulae Taurus and in Serpents and the emission nebulae of the type NGC 1499 and S48-S108. Later he convincingly proved (Refs. 40-44) the actual expansion of diffuse nebulae as well as the essential role of the inter-stellar magnetic field in this process. The existence of the correlation between the direction of the elongations of nebulae and the inter-stellar magnetic field disclosed by means of inter-stellar polarization of stellar light proves the validity of Shayn's hypothesis. /188
At present this interpretation of elongation nebulae of the indicated types cannot be doubted.

The case is somewhat different in the explanation of the fine fibrous structure of clouds of super novas: the magnetic field and the usual expansion of the hot substance cannot explain the existence of fibers with lengths a hundred times greater than their thickness. In 1946 Öort (Ref. 10) proposed the investigation of the thin fibrous system NGC 6960-6992-5 as a luminescent impact front which occurred during the motion of the expanding cloud in the inter-stellar substance. Pickelner (Ref. 45) calculated in detail the luminescence of this front and showed that the observations agreed quite well with the theoretical calculations if it is assumed that the thin fibers are formed during the intersection of two shock fronts. Moreover, the orientation of the fibers in the system NGC 6960-6992-5 relative to the inter-stellar magnetic field and the relation of the fibers in the Cancer nebula with the local magnetic field (see for example Ref. 46) attest to the major role played by the magnetic field in such objects. In the framework of magnetic hydrodynamics several models of the formation of such fine fibrous structures were proposed. Kaplan (Ref. 47) asserted that the magnetic field significantly arises in a shock wave and this results in induction fluxes. Under the influence of these fluxes the shock front dissipates into individual fibers. Pikel'ner and the author (Ref. 48) calculated a rough model of the "mechanical pinch"--the mechanism for the formation of fibrous structures as a result of turbulent motions which accompany the medium into the magnetic field. This mechanism has effective influence in the medium with a Reynolds magnetic number of greater than 100 to 200. In the ionized regions of intergalactic matter this condition is satisfied but the danger exists that diverse aspects of the hydromagnetic instability will hinder the development of the fibers. R. V. Polovin (Ref. 49) has

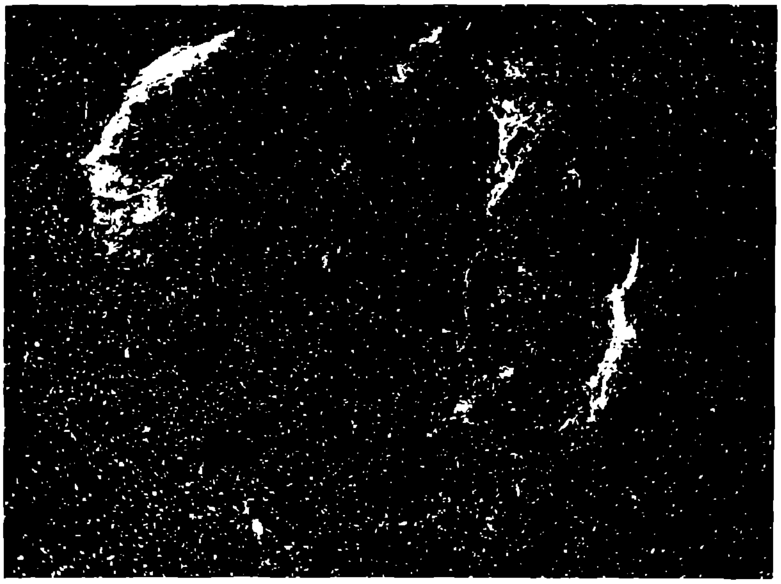


Fig. 3. Fine Fibrous Emission Nebula
NGC 6960-6992-5.

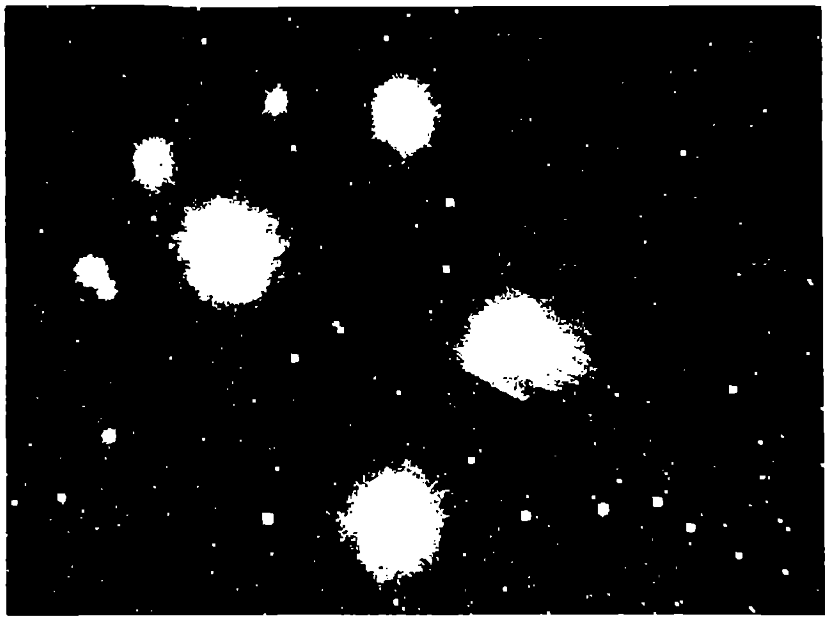


Fig. 4. Fine Fibrous Reflecting Nebula
in the Pleiades.

shown that in some cases the shock wave in the magneto-hydrodynamic medium splits up into two shock waves and even proposed the application of this effect for the explanation of the fine fibrous emission nebulae. Recently Pickelner (Ref. 50) found that during the expansion of the super nova clouds which contained a magnetic field, the pressure of the cosmic rays leads to the disintegration of these clouds into fibrous details. In the framework of this model the existence of nebular objects of two sharply distinct types in the source of radio Cas A can be simply explained.

As for the fine fibrous structures in the reflective nebulae (of the Pleiades type, they have not obtained at this time any physical interpretation. It can only be asserted (Ref. 51) that, in the first place, the filaments in such nebulae were always oriented along a local magnetic field and, in second place, the known physical processes, i.e., gravitational disintegration of a continuous medium, the multiplication of the density fluctuations in the magnetic field, the weak and strong /189 magneto-hydrodynamic distortions, "the mechanical" pinch-effect cannot explain the formation of these fine fibrous structures. Evidently, in a given case the limitation of our knowledge on the physical condition of matter in the regions H I to this time does not permit the construction of a satisfactory dynamic theory.

In conclusion of this survey it is necessary to emphasize that the investigation of the morphological peculiarities of the diffuse nebulae already has led to the occurrence of such important cosmogonical ideas as the concepts concerning the expansion and disintegration of the emission nebulae, the fundamental role of inter-stellar magnetic fields in the development of diffuse substance, the origin of emission nebulae and stars in a single process from other forms of matter, and also the explanation of many problems of cosmic gas-dynamics and cosmic magneto-hydrodynamics. All of this stimulates further research on the structure of diffuse nebulae for a deeper understanding of the physics and evolution of the inter-stellar medium.

REFERENCES

1. G. A. Schein and V. F. Gaze. Izvestiya Krymskoy Astrofizicheskoy Observatorii, 6, 3, 1950.

2. B. Stremgren. Astrofizicheskiy Sbornik (Astrophysical Symposium) Izd-vo Inostrannoy Literatury, 1949, p. 222.

3. B. A. Vorontsov-Velyaminov. Gazovye Tumannosti i Novye Zvezdy (The Gas Nebulae and Novae), Izd-vo A. N. SSSR, 1948.

4. S. Cederblad. Meddelanden Lunds Astron. Observ., Series II, No. 119, 1946.

5. E. E. Barnard. A Photographic Atlas of Selected Regions of the Milky Way, Washington, 1927.

6. J. E. Keeler. Publications of the Lick Observatory, 8, 1908.

7. I. Roberts. Photographs of Stars, Star Clusters and Nebulae, London, 1899.

8. J. C. Duncan. Astrophysical Journal, 53, 392, 1921; 57, 137, 1923; 63, 122, 1926.

9. O. Struve. Astrophys. J., 85, 194, 1937.

10. J. H. Öort. Monthly Notices of the Royal Astronomical Society, 106, 159, 1946.

11. A. D. Thackeray. Monthly Notices Roy. Astron. Soc., 110, 343, 1950.

12. D. A. Rozhkovskiy. Astron. Zh., 31, 318, 1954.

13. L. Spitzer. Astrophys. J., 120, 1, 1954.

14. E. A. Frieman. Astrophys. J., 120, 18, 1954.

15. D. Layver. Astrophys. J., 122, 1, 1955.

16. H. C. van de Hulst. Gas Dynamics of Cosmic Clouds, IAU Symposium No. 2, Amsterdam, 1955, p. 111.

17. D. Osterbrock. Astrophys. J., 125, 622, 1957.

18. B. J. Bok and E. F. Reilly. Astrophys. J. 105, 255, 1947.

19. V. G. Fesenkov and D. A. Rozhkovskiy. Astron. Zh., 29, 381, 1952.

REFERENCES (Cont'd)

20. S. Pottasch. Bulletin of the Astronomical Institutes of the Netherlands, 13, No. 471, 1956.

21. S. Pottasch. Bull. Astron. Inst. Neth., 14, No. 482, 1958.

22. F. D. Kahn. Bull Astron. Inst. Neth., No. 456, 1954.

23. F. D. Kahn. Kosmicheskaya Gazodinamika, Symposium (Cosmic Gas Dynamics) Izd-vo Inostrannoy Literatury, 1960, p. 245.

24. R. Ye. Gershberg. Astron. Zh. 38, 819, 1961.

25. R. Ye. Gershberg. Izv. Krymsk. Astrofiz. Observ., 26, 324, 1961.

26. R. Minkovski. Publications of the Astronomical Society of Pacific, /190 61, 151, 1949.

27. G. A. Schein and V. F. Gaze. Iav. Krymsk. Astrofiz. Observ., 7, 87, 1951.

28. G. A. Schein and V. F. Gaze. Izv. Krymsk. Astrofiz. Observ., 8, 80, 1952.

29. G. A. Schein and V. F. Gaze. Iav. Krymsk. Astrofiz. Observ., 10, 210, 1953.

30. G. A. Schein and V. F. Gaze. Astron. Zh., 30, 135, 1953.

31. G. A. Schein and V. F. Gaze. Astron. Zh., 31, 305, 1954.

32. G. A. Schein and V. F. Gaze. Izv. Krymsk. Astrofiz. Observ., 15, 3, 1955.

33. J. H. Öort. Bull. Astron. Inst. Neth., 12, No. 455, 1954.

34. J. H. Öort and L. Spitzer. Astrophys. J., 21, No. 6, 1955.

35. M. T. Savedoff and J. Greene. Astrophys. J., 122, 477, 1955.

36. V. I. Pronik. Izv. Krymsk. Astrofiz Observ., 23, 3, 1960.

37. R. Ye. Gershberg. Izv. Krymsk. Astrofiz. Observ., 25, 76, 1961.

38. F. E. Ross and N. R. Calvert. Atlas of the Milky Way, Chicago, 1934.

39. Ye. L. Ruskol. Astron. Zh., 27, 341, 1950.

REFERENCES (Cont'd)

40. G. A. Schein. Astron. Zh., 32, 110, 1955.

41. G. A. Schein. Astron. Zh., 32, 381, 1955.

42. G. A. Schein. Astron. Zh., 32, 489, 1955.

43. G. A. Schein. Astron. Zh., 33, 210, 1956.

44. G. A. Schein. Astron. Zh., 33, 305, 1956.

45. S. P. Pikelner. Izv. Krymsk. Astrofiz. Observ., 12, 93, 1954.

46. L. Woltjer. Bull. Astron. Inst. Neth., 14, No. 483, 1958.

47. S. A. Kaplan. Astron. Zh., 31, 358, 1954.

48. S. B. Pikelner and R. Ye. Gershberg. Astron. Zhur., 36, 785, 1959.

49. R. B. Polovin. Zhurnal Eksperimentalnoy i Teoreticheskoy Fiziki, 39, 1005, 1960.

50. S. B. Pikelner. Astron. Zh., 38, 21, 1961.

51. R. Ye. Gershberg. Izv. Krymsk. Astrofiz. Observ., 23, 21, 1960.

ELECTRON TEMPERATURE OF GASEOUS NEBULAE AND METHODS OF ITS DETERMINATION /191

By

V. I. Pronik

In the first part of this review, the works on the investigation of energy balance and the determination of the electron temperature in the gaseous nebulae are considered. In the second part, a review is given on the methods of the determination of the electron temperature from the observations.

The problem of investigating the electron temperature of gaseous nebulae is divided into three separate tasks:

1. The purely theoretical derivation of the electron temperature on the basis of a detailed analysis of the processes which occur in the gaseous nebulae, the processes which may in one way or another influence the energy of the free electrons of the nebula. To these belongs, for example, the works concerning the investigation of the energetic balance of free electrons.

2. The development of the methods of determination of the electron temperature by means of observations.

3. Derivation of observation data and determination of electron temperature by the developed method.

In the present article the state of the problem is examined according to the first two problems.

I. The Energetic Balance of Free Electrons and Electronic Temperature of Gaseous Nebulae

Although the first observations of the spectra of nebulae were concluded basically before 1918 (Ref. 1) the information concerning their electron temperature T_e appeared only in 1939 when Ambartsumyan (Ref. 2) developed for the first time a method for the determination of T_e according to the lines $N_1 + N_2$ and λ 4363 of a twice ionized oxygen.

Disregarding the impacts of the second type Ambartsumyan obtained the formula for the relation of the intensity of the lines $N_1 + N_2$ and λ 4363

$$\frac{E_{N_1+N_2}}{E_{\lambda 4364}} = \frac{v_{1\to 2}}{v_{2\to 3}} \left[1 + \frac{A_{3\to 1} + A_{3\to 2}}{A_{3\to 2}} \frac{\sigma_{1\to 2}}{\sigma_{1\to 3}} e^{\frac{x_{2-3}}{kT_e}}\right], \qquad (1)$$

where the conditions ^3P, ^1D and ^1S correspond to 1, 2, and 3. Here A denotes the probability of the spontaneous transition and σ, the effective cross section of excitation. Roughly considering that $\sigma_{1\to 2}/\sigma_{1\to 3} = 1$ (the values of the effective cross sections were at that time still not known) and accepting according to (Ref. 2) the mean value of the relation $E_{N_1+N_2}/E_{4363}$ as being equal to 30, Ambartsumyan obtained from formula (1) the average value of the electron temperature of a nebulae $T_e \approx 7000°C$.

At approximately the same time Baker, Menzel and Aller (Ref. 3) on the basis of the equation for the radiant equilibrium* and the condition of the stationarity of ionization in the nebula obtained a relation between the electron temperature of the gas and the temperature of the radiation of the star T_* beyond the range of the Layman series for a case of a purely hydrogen nebula. For an optically fine nebula ($\tau_{L_c} < 1$) this relation has the following form:

$$\frac{AT_* \displaystyle\int_{y_1}^{\infty} \frac{dy}{e^y - 1}}{\displaystyle\int_{y_1}^{\infty} \frac{dy}{y\,(e^y - 1)}} = \frac{A^2T_e^2 + aAT_e + \frac{1}{2}\, G_{T_e} - \frac{1}{2} \sum_1^{\infty} \frac{S_n}{n^5}}{G_{T_e}}, \qquad (2)$$

---

*This equation is a description of the physical condition under which the quantity of energy absorbed by the atoms should be equal to the quantity of radiated energy for the units of volume and time.

where

$$A = \frac{k}{2hRZ^2}\ ; \quad y = \frac{h\nu}{kT_*}\ ; \quad y_1 = \frac{k\nu_\infty}{kT_*}\ ; \quad a = \sum_1^\infty \frac{1}{n^3}\ ;$$

$$S_n = e^{\frac{hRZ^2}{n^2kT_e}} \left[- E_i \left(- \frac{hRZ^2}{n^2kT_e} \right) \right] ,$$

$$GT_e = \sum_1^\infty \frac{S_n}{n^3} .$$

Symbolically, equation (2) can be written in the form

$$f_1\ (T_*) = f_2\ (T_e), \tag{3}$$

where the left term is dependent only on T_* and the right term is the function dependent on only T_e. The relation which they obtained between the temperature of the star and the temperature of the gas for the optically fine nebula of pure hydrogen is shown in Table 1. /193

Table 1. The Relation of T_* and T_e

T_*	T_e	T_*	T_e
5 000	5 000	80 000	57 000
10 000	9 500	160 000	92 000
20 000	18 000	320 000	132 000
40 000	34 000		

Later Menzel, Aller, and Hebb (Refs. 4, 5) calculated the effective cross sections of the excitations by electron impact of a D and S level of the ion O III and by utilizing the intensity of the lines $N_1 + N_2$ and λ 4363 which they observed, they calculated the values of T_e for twenty or thirty planetary nebulae by the Ambartsumyan method.

Thereby, they took into consideration also the influence of the impacts of the second type. As a result the formula (1) took the form

$$\frac{E_{N_1+N_2}}{E_{\lambda 4363}} = 4.07\ e^{33000/T_e}. \tag{4}$$

For an overwhelming majority of nebulae the electron temperature proved to be in the limits of 7,000 to 10,000°C. The temperature did not show any dependence on the degree of excitation of the nebula and on the temperature of the central star which was found by the Zanster method.

The divergence between the observed electron temperature and the theoretical value of it obtained for a purely hydrogen nebula (Ref. 3) was soon explained by Menzel and Aller (Ref. 6) as the effect of the cooling off of the nebula on the ions O III. They showed that if we consider the radiation of the nebula in the prohibited lines (O III) in the equation of the radiant equilibrium, then in the right term of equation (2) an additional member will appear, equal to

$$C\ (T_e) = 1.63\ \frac{T_e}{G_{T_e}}\ e^{-\frac{x_{PD}}{kT_e}}\ \frac{N_p}{N_i}\ (1 - \beta). \tag{5}$$

Here N_p is the number of atoms of O III in the basic state; N_i is the /194
number of protons. The factor $(1 - \beta)$ takes into consideration the impacts of the second type

$$\beta = \frac{8.54 \cdot 10^{-6}\ \frac{N_D N_e}{\omega_D}\ \frac{\Omega_{DP}}{T_e^{1/2}}}{8.54 \cdot 10^{-6}\ \frac{N_D N_e}{\omega_D}\ \frac{\Omega_{DP}}{T_e^{1/2}} + N_D A_{D\to P}} = \frac{1}{1 + 5.85 \cdot 10^{5}\ \frac{A_{D\to P}}{\Omega_{P\to D}}\ \frac{T_e^{1/2}}{N_e}}\ . \tag{6}$$

The atomic parameter Ω is connected with the effective diameter cross section σ by the correlation

$$\Omega = \frac{1}{(2J + 1)} \frac{h^2}{4\pi m^2} \frac{\sigma}{v^2}, \tag{7}$$

where 2J + 1 is the statistical weight of the lower state and v is the velocity of the electron.

In the symbolic designations of equation (3) this new equation of the radiant equilibrium is written thus:

$$f_1 (T_*) = f_2 (T_e) + C (T_e). \tag{8}$$

The corrective term $C (T_e)$ proves to be extremely essential. For example there is enough relative abundance of ions O III on the order of $N_{O\ III}/N_i \approx 10^{-4}$ to lower the electron temperature of the nebula from 57,000 to 8,000°C and the temperature of the exciting star of 80,000°C. Recently Aller (Ref. 7) essentially improved this method for determining electron temperature by presenting the energy of the nebula being radiated in the prohibited lines as follows

$$E_i = E (H_\beta) \sum_i \frac{I_i}{I_{H_\beta}}, \tag{9}$$

where E_{H_β} is the energy emitted in the line H_β and $\sum_i I_i/I_{H_\beta}$ is the observed relation of the total intensity of all the prohibited lines to the intensity of the line H_β. Thereby the corrective term in equation (8) takes on the form

$$C (T_e) = 1.61 \cdot 10^{-3} \frac{b_4 (T_e)}{G_{T_e}} e^{\frac{h\nu_\beta}{kT_e}} \sum_i \frac{I_i}{I_{H_\beta}}, \tag{10}$$

where $b_4 (T_e)$ characterize the degree of deviation of the fourth level population hydrogen atom from its population during thermal dynamic equilibrium (Refs. 8, 9). By using the values of T_* found spectrophotometrically or by the Zanster method and the observed magnitude /195 $\sum_i I_i/I_{H_\beta}$ Aller (Ref. 7) found in this way the temperature of a series

of nebulae. These values of T_e as a rule are essentially greater than the temperatures found according to the relation $E_{N_1+N_2}/E_{\lambda\,4363}$.

In 1941 V. V. Sobolev (Ref. 10) proposed a method of determining T_e, based on the investigation of the energetic balance of electron gas. Thereby, the energy losses of free electrons on excitation and ionization of hydrogen atoms by an electron impact were taken into consideration for the first time. The essence of this method is contained in the fact that the quantity of energy which electrons obtain during the photo-ionization of hydrogen should be equal to the full amount of energy that they lose during interaction with other atoms and ions plus the energy of recombining electrons. Since the latter depend on the electron temperature of the nebula, the condition of equality of gain and loss of energy per units of volume and time, determines in itself the sought value of the equilibrial electron temperature of the nebula. Considering that electrons lose energy in three ways: by radiation in a continuous spectrum (free transitions) on excitation of the lines $N_1 + N_2$, and in the course of inelastic collisions with hydrogen atoms, Sobolev obtained on the basis of the energy conservation law, an equation in the following form

$$AT_* = BT_e + C\,\frac{I_{4959}}{I_{H_\beta}} + D\,\frac{N_{H_1}}{N_{H^+}}, \tag{11}$$

where I_{4959}/I_{H_β} is the relation of intensity of the line 4959 (O III) to the intensity of H_β, N_{H_1}/N_{H^+}, i.e., the relation of the number of neutral hydrogen atoms to the number of protons. When $\tau_{L_c} < 1$ the coefficient A has the following value:

$$A = \frac{\displaystyle\int_{x_0}^{\infty} \frac{dx}{e^x - 1}}{\displaystyle\int_{x_0}^{\infty} \frac{dx}{x\,(e^x - 1)}} - x_0; \qquad x = \frac{h\nu}{kT_*} \qquad x_0 = \frac{h\nu_\infty}{kT_*} \tag{12}$$

and consequently depends only on the temperature of the star. The coefficients B, C and D are the function of the nebula's electron temperature only. It is easily noticed, that, in principle, the equations (8) and (11) are identical. Only equation (11) has an additional term which takes into consideration the excitation of hydrogen by an electron impact. The electron temperature of a series of planetary nebulae found by means of equation (11) is within the limits of 13 to 19,000°C, i.e., it is an identical temperature to the one obtained by Aller (Ref. 7) considerably later with a similar method. The loss of /196 energy by electrons during inelastic collisions with neutral hydrogen atoms, constitutes from 15 to 60% of the total electron energy according to (Ref. 10).

Along with the works by Menzel, Aller and Sobolev it is also necessary to mention the works by Spitzer (Refs. 11-14) on the determination of the electron temperature of the H I and H II regions of inter-stellar gas. As in the works of preceding authors, the works by Spitzer are of a theoretical character. During 1947 to 1949 Spitzer investigated (Refs. 11-12) all possible processes of heating and cooling of gas, which occurs in the regions H II near morning stars and in the regions of neutral hydrogen with the purpose of obtaining a theoretical determination of electron temperature of inter-stellar gas. In the regions H II, in which we are interested in the given case, these processes are reduced to the one which has already been examined earlier by Sobolev and Aller, with the exception of processes related to the presence of dust. Generally speaking, the existence of dust in the hot region of ionized hydrogen is not at all understandable, but the classical example of the Major Nebula in Orion shows that this is evidently the case. According to (Ref. 12) the presence of dust particles if they are dielectric, leads to the adsorption of free electrons. Such adsorption is related to the condensation of electrons on the surface of the dust particles and their subsequent breaking off as a result of the impacts by free electrons. The electrons which cling to the surface of the dust particle can also be neutralized by means of recombination with positive ions. However, the probability that a dust particle will encounter an ion ($\xi_i \sim 1 \div 10^{-4}$) is on the average less than the probability of a collision with an electron ($\xi_e \sim 0.5 \div 0.1$). If we disregard completely the neutralization, considering roughly that the number of adherences of the electrons is accurately equilibrial to the number of impacts during which the free electron dislodges adhering electrons from the surface of the dust particle, then the over-all energy being lost by electrons as a result of this process in volume and time units is equal to

$$N_e N_g 3 \left(\frac{2\pi}{m_e}\right)^{1/2} \xi_e \, (kT_e)^{3/2} \, a^2 e^{\frac{eV}{kT_e}} , \qquad (13)$$

where N_g is the number of dust particles per unit volume, a is their average radius, V is the electrostatic potential of the dust particle. The latter according to van de Hulst (Ref. 15) depends on T_e and is equal to $2.2 \frac{T_e}{10{,}000^\circ C}$ v.

The adsorption of electrons on the dust particles becomes essential when $N_g/N_H > 10^{-11}$ (the average relative abundance of dust for the interstellar medium considered equal to $N_g/N_H = 10^{-12}$). The values of /197 the equilibrial temperature in the regions H II near O-B-stars obtained by Spitzer and Savedov (Ref. 13) on the basis of the law of preservation of energy are given in Table 2. The values of T_e were calculated for different densities ($N_H = 10^{-2}$, 1 and 10^2 cm^{-3}) and for a diverse abundance of cooling ions ($N_i/N_H = 2 \cdot 10^{-3}$ and $2 \cdot 10^{-4}$). The abundance of dust particles was taken to be standard.

Table 2. The Value of Equilibrial Temperature

N_H	N_i/N_H	T_e	
		Near O-star	Near B-star
10^{-2}	$2 \cdot 10^{-3}$	6 800°	4 800°
	$2 \cdot 10^{-4}$	12 000	9 500
1	$2 \cdot 10^{-3}$	6 900	5 000
	$2 \cdot 10^{-4}$	12 000	9 500
10^2	$2 \cdot 10^{-3}$	7 600	5 700
	$2 \cdot 10^{-4}$	13 000	10 000

As should have been anticipated within these limits of densities the electron temperature is almost independent of the density of the nebula but is extremely essentially dependent on the relative abundance of cooling ions and the spectral class of the exciting star.

Somewhat later Spitzer (Ref. 14) found the average equilibrium temperature of the regions H II by the same method equal to approximately 8000°C. Thereby, a certain mean chemical composition (N_i/N_H) and a standard content of the dust were adapted.

The works of Menzel, Aller, Sobolev and Spitzer which were discussed above are of a theoretical nature and the values of T_e which they yielded are also most probably theoretical. The method which they are using is the same energy balance method. The equation of energy balance is of interest when the electron temperature of the nebula is known, since, in this case, we can either define more accurately the data concerning the radiation of the star beyond the Layman limit or confirm the accuracy of our concepts concerning the processes which are responsible for the cooling and heating of free electrons in the nebulae. In view of the great importance of this method it is extremely desirable to define more accurately and in greater detail examinations of the energy balance of the electron gas.

The present state of astrophysics permits a more precise definition of the energy balance of the nebula. We have in mind the following /198 three moments: 1) the consideration of more accurate data concerning the radiation of stars beyond the Layman limit (until now this was considered to be a Planck radiation with a temperature equal to the effective temperature of the star), 2) the utilization of the latter more accurate data concerning the chemical composition of the nebulae and also consideration of the change in ionization of the nebula with distance from the exciting star, and finally, 3) the utilization of more accurate values for the probability of the excitation of hydrogen by an electron impact.

The energy distribution in the spectra of hot stars in the region of the wavelengths $\lambda < 9.2$ Å has been obtained by many authors [Underhill (Refs. 16-18), Pecker (Ref. 19), Traving (Refs. 20-21), Saito and Uesugi (Ref. 22)] in their calculations of a model of the atmosphere of these stars. As was to be anticipated such a distribution differs from Planck's law, primarily because of the presence of skips beyond the limits of the main series. When $T_* \geq 50,000°C$ the skips near the Layman limit smooth out and the radiation of such stars in the region of 300 Å $< \lambda <$ 912 Å is already close to the black body radiation. By using these data concerning the distribution of energy in the spectra of hot stars beyond the Layman limit we have calculated the averate energy value of the photo-ionized electron for an optically fine* purely hydrogen nebula (Table 3).

---

*The condition of transparency for L_c quanta has been taken directly from the definition of the Stremgren zone (Ref. 23) according to which the inequality $\tau_{L_c} < 1$ is valid for practically the whole volume of the nebula (of the zones H III), with the exception of a very fine bordering layer where $\tau_{L_c} \gtrsim 1$.

Table 3. Average Value of Energy of the Photo-Ionized Electron

Spectral Class of Exciting Star	$\epsilon_0 \cdot 10^{12}$ erg	Author of Model
WR ($T_* = 80,000°$)	7.4	Underhill [18]
WR ($T_* = 62,000°$)	7.4	" "
WR ($T_* = 50,000°$)	7.4	" "
05 V	4.8	Underhill [16]
07.5 V	4.7	Saito and Uesugi [22]
09 V (τ Sco)	4.2	Traving [20]
09.5 V	4.2	Underhill [17]
BO V(10 Lac)	4.2	Traving [21]
BO.5 V	3.4	Saito and Uesugi [22]
B1 V	3.6	Pecker [19]
B1.5 B	3.6	" "

In actual nebulae due to the presence of an admixture of helium/199 atoms and other elements, the average energy obtained by electrons in the ionization of hydrogen and helium will differ from the value of $\overline{\epsilon}_0$ for a purely hydrogen nebula primarily for two reasons:

1) During the ionization of the atoms He I and He II new electrons appear (free electrons of non-hydrogen origin) initial energy of which $(h\overline{\nu} - x_0)$ can differ from the average energy of photo-electrons of a hydrogen origin and dependent on the type of distinction (greater or smaller) the over-all energy which is ascribed to one electron is correspondingly augmented or diminished;

2) The second reason is related to a change in spectral composition of radiation which ionizes the hydrogen as a result of the absorption by atoms of He I and a radiation with $\lambda < 504$ Å.

Taking all of this into consideration, we have calculated the average energy which is obtained by free electrons during the ionization of hydrogen and helium in various zones of the nebula. These zones are characterized by the diverse degree of helium ionization (Table 4).

Table 4. Mean Energy of Free Electrons Obtained During Ionization of Hydrogen and Helium

Spectral Class of Exciting Star	$\epsilon_0 \cdot 10^{12}$ erg		
	In Zone He I	In Zone He II	In Zone He III
WR	5.3	7.9	9.7
WR	5.2	7.9	7.7
WR	5.1	7.9	6.7
O 5 V	4.8	5.3	-
O7.5 V	4.7	5.0	-
O 9 V	4.2	-	-
O9.5 V	4.2	-	-
BO V	4.2	-	-
BO.5 V	3.4	-	-
B1 V	3.6	-	-
B1.5 V	3.6	-	-

The condition of energy balance signifies that the energy of the recombining electron plus all energy losses by the electron during its stay in the free state should be equal to the initial energy of the photo-electron

$$\bar{\epsilon}_0 = \bar{\epsilon}_{rec} + \epsilon_{loss} \ . \tag{14}$$

The energy losses occur as a result of the excitation by an electron/200 of the prohibited lines of various ions, the excitation of hydrogen by the electron impact and also as a result of the "free" transitions $\left(\epsilon_{loss} = \sum_i \epsilon_i + \epsilon_H + \epsilon_{f-f}\right)$.

The energy which is lost by the electron on the excitation of the given prohibited line during the entire period of its stay in a free state is equal to

$$\epsilon_i = 8.54 \cdot 10^{-6} \frac{N_i}{N_H} \frac{\Omega\theta\,(T_e)}{\tilde{\omega}_A} \frac{e^{-\frac{x_{AB}}{kT_e}}}{T_e^{1/2} \sum_1^{\infty} C_n\,(T_e)} k\nu_{AB}\,, \tag{15}$$

where N_i/N_H is the qualitative abundance of ions of the given type and the function $\theta\,(T_e)$ is the corrective factor which expresses the dependence of the parameter Ω on the temperature (Ref. 24). Symbolically, the expression (15) can be written

$$\epsilon_i = \frac{N_i}{N_H} \psi\,(T_e), \tag{16}$$

where $\psi\,(T_e)$ for a given line is a function of the temperature alone. Then it is apparent that the over-all energy which is lost by the electron and the excitation of the lines of various ions

$$\sum_i \epsilon_i = \sum_i \frac{N_i}{N_H} \psi_i\,(T_e). \tag{17}$$

In order to calculate $\sum_i \epsilon_i$ it is necessary to know not only the chemical composition of the nebula, but also the change of ionization of the different elements with the change in distance from the exciting star.

Table 5. The Mean Chemical Contents of the Nebulae

Element	Abundance
H_e	0.15
N	$1.6 \cdot 10^{-4}$
O	$6 \cdot 10^{-4}$
S	$1 \cdot 10^{-4}$

In Table 5 is shown the mean chemical composition of the nebulae which is obtained according to the observation data of various authors (Refs. 25-30) (the abundance of hydrogen atoms is accepted to be on the order of unity) and Table 6 gives a presentation concerning the ionization of sulfur, nitrogen, and oxygen in various zones of the ionization of helium. As a basis for the construction of Table 6 the observed distribution of atoms O, N, and S was accepted according to their stages of ionization in the nebula Orion (Ref. 26) and NGC 7027 (Ref. 25). The numerical values give the approximate content of ions in percentages from the complete number of atoms of the given element.

Table 6. Ionization of Oxygen, Nitrogen, and Sulfur in Various Zones of a Nebula /201

Ion	Zone			
	He I	He $II_{(1)}$	He $II_{(2)}$	He III
[O II]*	100	70	30	0
[O III]	0	30	60	30
O IV	0	0	10	70
[S II]	70	30	10	0
[S III]	30	50	60	10
S IV	0	20	30	60
[N II]	90	50	25	0
[N III]	10	50	67	10
N IV	0	0	8	70

*Within the square brackets are enclosed ions, which participate in the disorbtion of nebula.

The energy which is lost by the electron during the inelastic collisions with atoms of hydrogen is equal to

$$\epsilon_H = \frac{N_{H_1}}{N_{H^+}} D\ (T_e), \tag{18}$$

where

$$D\ (T_e) = \frac{\sum_{1}^{\infty} b_{1n}x_{1n} + \gamma_k x_0}{\sum_{1}^{\infty} C_n\ (T_e)} .$$

The probabilities of excitation and ionization of hydrogen by the electron impacts are adopted in accordance with the work by Chamberlain (Ref. 31). The function $D\ (T_e)$ grows very rapidly with the increase in temperature; therefore, the existence of such a mechanism of electron adsorption as, for example, inelastic collision with neutral hydrogen atoms does not make it possible to raise the electron temperature of the nebula essentially greater than 20,000°C.

The energy which is lost by the electron during the "free-free" /202 transitions during the entire period of its presence in the free state is equal

$$\epsilon_{f-f} = 1.42 \cdot 10^{-27} \frac{T_e^{1/2}}{\sum_{1}^{\infty} C_n\ (T_e)} . \tag{19}$$

As a result of the fact that the probability of the recombination of the electron is dependent on its velocity (the greater the velocity of recombination the less the velocity of the electron) the average energy of the recombining electron is about half as great as the mean energy of the free electron ($3/2\ kT_e$). An exact calculation indicates that the relationship

$$\frac{\bar{\epsilon}_{rec}}{\frac{3}{2}\ kT_e} = 0.52 \tag{20}$$

is practically constant in the limits of temperature from 3,000 to 30,000°C.

By knowing the initial energy of the photo-electron and the energy loss by the electron at various temperatures, we can find from

equation (14) the value of equilibrial electron temperature in each zone of the nebula. The results of the calculations are shown in Table 7.

This table indicates the dependence of the electron temperature of the nebula on the spectral class of the exciting star and its change in limits of one and the same nebula. This change is stipulated by the change in ionization in the nebula and also by the change in mean energy of the ionizing quantum.

Table 7. Equilibrial Electron Temperature in Various Zones of the Nebula

Spectral Class	He I	He II_1	He II_2	He III
WR	8 000°	9 700°	10 300°	14 000°-16 000°
05-07.5	7 600	8 400	8 800	-
09-BO	7 000	-	-	-

II. THE METHODS OF THE DETERMINATION OF THE ELECTRON TEMPERATURE OF GASEOUS NEBULAE FROM OBSERVATIONS /203

1. The Determination of the Electron Temperature by the Relation of Intensity of Auroral and Nebular Lines

The fundamental and most accurate method for the determination of T_e of gaseous nebulae is the method based on a comparison of intensities of two prohibited lines of one and the same ion which are excited by electron impacts of the principal level. However, the potentials of excitation of the lines relative to this basic level are distinct (Fig. 1). The greater that the difference is in the potentials of excitation of the upper levels of these lines the less is the inaccuracy T_e in the determinations of the observed values of the line intensities. The energy which is radiated by a unit of volume in the nebula in the prohibited line is equal to

$$E_{AB} = 8.54 \cdot 10^{-6} \frac{N_A N_e}{\tilde{\omega}_A T_e^{1/2}} \times$$

$$\times \Omega_{AB} \theta (T_e) e^{-\frac{x_{AB}}{kT_e}} h\nu_{B\to A} . \tag{21}$$

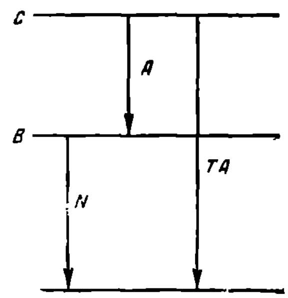


Fig. 1. Diagram of the Lower Levels of an Atom. The arrows indicate the transitions during which the nebular (N), auroral (A) and transauroral (TA) lines are radiated.

Thereby, it is assumed that each excited atom is lowered spontaneously emitting a quantum with an energy $h\nu_{B\to A}$. Generally taking into consideration impacts of the second type, this formula has the following form (Ref. 32):

$$E_{AB} = N_A \frac{\tilde{\omega}_B}{\tilde{\omega}_A} e^{-\frac{x_{AB}}{kT_e}} \left(\frac{A_{B\to A} \cdot \alpha_{BA} \cdot N_e}{A_{B\to A} + \alpha_{BA} \cdot N_e} \right) h\nu_{AB}, \tag{22}$$

where $A_{B\to A}$ is the probability of the spontaneous transition and

$$\alpha_{BA} = 8.54 \cdot 10^{-6} \frac{\Omega_{AB}\theta\ (T_e)}{\tilde{\omega}_A \cdot T_e^{1/2}} \ . \tag{23}$$

A critical value of density N_e^k usually introduced is determined by the correlation

$$N_e^k = \frac{A_{B\to A}}{\alpha_{BA}} \ . \tag{24}$$

From expressions (22) and (24) it is obvious that if $N_e \ll N_e^k$ then the /204 number of radiation acts is not dependent on $A_{B\to A}$ and we have the case described by formula (21). On the other hand, if $N_e \gg N_e^k$ then the amount of radiation is not dependent on N_e but is proportional to the quantity of radiating atoms and the probability of spontaneous transition:

$$E_{AB} = N_A \cdot \frac{\tilde{\omega}_B}{\tilde{\omega}_A} e^{-\frac{x_{BA}}{kT_e}} \cdot A_{B\to A} \cdot h\nu_{BA}, \tag{25}$$

that is, we have a case of Boltzmann distribution.

In the formula which expresses the relation of the intensity of the auroral and nebular lines the temperature relation apparently will be represented primarily by the following exponential factor

$$\frac{I_A}{I_N} \approx e^{-\frac{x_{13}}{kT_e} + \frac{x_{12}}{kT_e}} = e^{-\frac{x_{23}}{kT_e}} \ . \tag{26}$$

If for such a pair of lines, apart from the potentials of excitation, the probabilities of spontaneous transitions are essentially distinct then with densities close to critical or greater, the relationship I_A/I_N (or I_{TA}/I_N), in addition to the temperature, will still depend on electron density of the nebula. In this case it is necessary to have two such correlations (let us say for two types of ions). The concurrent solution of two equations yields the value of N_e and T_e. It is easy to

consider that in a case where potentials of excitation of the upper levels of two lines are equal or very close and the probability of spontaneous transitions are excellent, then when $N_e \approx N_e^k$ the relation I_A/I_N is dependent only on the electron density of the nebula (the method of determination of the electron density according to relation I_{3726}/I_{3729} was proposed by Osterbrock and Seaton (Ref. 32)).

The method for the determination of T_e with respect to the lines $I_{4363}/I_{N_1+N_2}$ of the ion [O III] is the most extensive and is based on the principle described earlier. If the equation for the statistical equilibrium for the upper two levels $^1D_2$ and $^1S_0$ of the ion O III (we shall designate them by 2 and 3) is written down and the population of the main level is excluded from the ion, then we shall obtain the relation of the populations of the levels $^1S_0$ and $^1D_2$:

$$\frac{N_{^1S_0}}{N_{^1D_2}} = \frac{\frac{\tilde{\omega}_3}{\tilde{\omega}_2} e^{-y} \left\{ 1 + \frac{\Omega_{23} e^{-y}}{\Omega_{21}\theta} + \frac{A_{21}\tilde{\omega}_2}{\Omega_{12}\theta \cdot N_e c} \right\}}{1 + \frac{A_{32} + A_{31}}{\Omega_{31}} \frac{\tilde{\omega}_3}{\theta N_e C} + \frac{\Omega_{31} e^{-y}}{\Omega_{21}\theta} + \frac{A_{32}\tilde{\omega}_3 e^{-y}}{\Omega_{12} N_e \cdot C\theta}}, \tag{27}$$

where /205

$$y = \frac{x_{13} - x_{12}}{kT_e} ; \quad C = \frac{8.54 \cdot 10^{-6}}{T_e^{1/2}} ; \quad \theta = 1 + \frac{\Omega_{32}}{\Omega_{31}} .$$

By knowing the probabilities of spontaneous transitions for the lines λ 4363 and $N_1 + N_2$ it is possible to find also the relation of the intensities of these lines. By using the most accurate values of the parameters Ω (Ref. 33) and the probabilities of spontaneous transitions (Ref. 34), we will obtain for the relationship $I_{4363}/I_{N_1+N_2}$ the expression (Ref. 26)

$$\frac{I_{4363}}{I_{N_1+N_2}} \sim 13.1 \cdot 10^{-\frac{14300}{T_e}} \left\{ \frac{1.01 + 2640 \dfrac{T_e^{1/2}}{N_e}}{1.01 + 2.52 \cdot 10^5 \dfrac{T_e^{1/2}}{N_e}} \right\}. \qquad (28)$$

It is obvious from this formula that in the case when the densities $N_e \ll 10^5$

$$\frac{I_{4363}}{I_{N_1+N_2}} \approx 0.137 \cdot 10^{-\frac{14300}{T_e}} . \qquad (29)$$

In the case of great densities ($N_e > 10^7$ star-like planetary nebulae, clouds of symbiotic stars of retarded spectral classes, etc.)

$$\frac{I_{4363}}{I_{N_1+N_2}} \approx 13.1 \cdot 10^{-\frac{14300}{T_e}} . \qquad (30)$$

Thus, in both boundary cases the relation $I_{4363}/I_{N_1+N_2}$ is not dependent on the density and only when $10^4 < N_e < 10^8$ is it a function of the temperature and density. For the relation of intensities of analogous lines of the ion [O II] Seaton and Osterbrock (Ref. 32) obtained the following expression

$$\frac{I_{7320+30}}{I_{3727}} = 0.18 \cdot e^{-y} \times$$

$$\times \left\{ \frac{1 + 0.138\, t\, (1 + 0.38\, e^{-y}) + 38.4 \cdot 10^{-4}\, t^2 \left(1 + 0.78 e^{-y} + 0.15\, e^{-\frac{17000}{T_e}}\right)}{1 + 0.36 e^{-y} + 5.3 \cdot 10^{-2}\, t \left(1 + 0.82 e^{-y} + 0.5 e^{-\frac{17000}{T_e}}\right)} \right\},$$

where

$$t = N_e/T_e^{1/2} \text{ and } y = \frac{8500}{T_e} .$$

This correlation is valid only when $N_e < 10^6$. From formula (31) it /206 is obvious that the relation $I_{7320+30}/I_{3727}$ is dependent on the electron density when $N_e > 10^2$. And only at very low densities is it dependent on the density and is equal to

$$\frac{I_{7320+30}}{I_{3727}} = \frac{0.18 \cdot e^{-\frac{8500}{T_e}}}{1 + 0.35 \cdot e^{-\frac{8500}{T_e}}} . \qquad (32)$$

Quite often, especially when the density of the nebula is given greater than 10^3 to 10^4, for the determination of N_e and T_e a system of equations is used, one of which is equation (28) and another is an analogous equation which is written for the line of the once-ionized nitrogen:

$$\frac{I_{5755}}{I_{6548+84}} = 61.5 \cdot 10^{-\frac{10820}{T_e}} \left\{ \frac{1.01 + 320 \frac{T_e^{1/2}}{N_e}}{1 + 19.4 \cdot 10^4 \frac{T_e^{1/2}}{N_e}} \right\} . \qquad (33)$$

The concurrent solutions of equations (28) and (33) are often produced by a graphic method: the relations between N_e and T_e with the data of the observed relations $I_{4363}/I_{N_1+N_2}$ and $I_{5755}/I_{6548+84}$ are given in the form of curves, the point of intersection of which determines the unknown values of N_e and T_e. This method of the determination of temperature and density at first was applied by Seaton (Ref. 35). At small densities the relation $I_{5755}/I_{6548+84}$ is equal to

$$\frac{I_{5755}}{I_{6548+84}} = 0.1 \cdot 10^{-\frac{10820}{T_e}} . \qquad (34)$$

Therefore, it can serve as an independent method for the determination of T_e on the level with equation (28).

Seaton (Ref. 35) proposed other equations which relate electron temperature and density of the nebula. One of them is based on the supposition that

$$\frac{N_1 \text{ (O I)}}{N_1 \text{ (N I)}} = \frac{N_1 \text{ (O II)}}{N_1 \text{ (N II)}} , \qquad (35)$$

where N_1 is the number of atoms in a given ion in the primary state. The validity of the equality (35) is argued by the fact that the ionization potentials and the cross section of photo-ionization for the ions O I and N I are very close. By expressing N_1 in terms of intensity of nebular lines of the respective ions, instead of (35) we obtain /207

$$\frac{I_{[\text{O II}]} \cdot I_{[\text{N II}]}}{I_{[\text{N I}]} \cdot I_{[\text{O II}]}} = 0.146 \frac{(1 + 3.4\ t_e^{1/2} x)\ (1 + 1.6\ x)\ e^{\frac{2.16}{t_e}}}{(1 + 0.1\ x)} \qquad (36)$$

where

$$x = 10^{-4} \frac{N_e}{t_e^{1/2}} , \quad t_e = 10^{-4}\ T_e .$$

Analogous equations can be obtained also on the basis of other suppositions of the type (35). Böhm (Ref. 36), for example, considers that

$$\frac{N_1 \text{ (O II)}}{N \text{ (O)}} = \frac{N_1 \text{ (S II)}}{N \text{ (S)}} \text{ or } \frac{N_1 \text{ (O III)}}{N \text{ (O)}} = \frac{N_1 \text{ (Ne III)}}{N \text{ (Ne)}} , \qquad (37)$$

where N (O), N (S), and N (Ne) are the total quantities of the atoms of the given elements. The relationships N (O)/N (S) and N (O)/N (Ne) are considered known and constant in all nebulae. However, the equalities of

the type (37) are a very rough approximation; they cannot be substantiated since neither the coefficients of absorption for the ions nor the field of radiation inside the nebula are known with any accuracy. As far as concerns equations of the type (28), (31), (33) which express the relationships of the intensities of nebular and auroral lines they can be written for many other ions, for example [S II], [S III], [O I], [N I]. However, in practice, as the method for determining T_e they cannot be utilized because of the weak intensity of the lines, which enters into these relations.

2. The Determination of T_e Through a Continuous Spectrum of the Nebula

The possibility of the determination of electron temperature, through a continuous spectrum of the nebula beyond the limits of the Balmer series, is connected with the fact that the intensity of continuous radiation in this region of the spectrum is dependent on T_e, as

$$e^{-\frac{h(\nu - \nu_n)}{kT_e}}, \tag{38}$$

where ν_n is the frequency which corresponds to the limit of the Balmer series. For the determination of T_e with this method, it is sufficient to measure the intensity of the Balmer continuum in two different frequencies. The method is very sensitive to mistakes in observations. It is very difficult to avoid such mistakes since the observations are conducted in the ultraviolet region of the spectrum. It is sufficient to mention that the first attempts to determine T_e by this method were unsuccessful (Refs. 37, 38).

Sealley's theory yielded one more method of determination of temperature which is based on the magnitude of the Balmer skip /208

$$D = \lg \frac{I_{\lambda<3686\ A}}{I_{\lambda>3686\ A}}. \tag{39}$$

By utilizing the theoretical formulas, Barbier (Ref. 39) calculated the magnitude of the Balmer skip during different electron temperatures considering that the continuous radiation of the nebula is postulated only by the freely-bound and free-free transitions. However, the

theory and observations indicated a significant divergence. Later with the development of the theory of the two-quantum continuous radiation (Refs. 40, 41) the conformity was improved between the theoretical Balmer skip and the observed skip. As Andrillat showed (Ref. 42) the observed magnitude of the Balmer skip is in good conformity with its theoretical value calculated by Seaton (Ref. 43) at a temperature close to T_e which was found along the lines [O III]. Thereby, 2 s → 2p transitions, which were brought about by electron collisions, were also taken into consideration.

The determination of the electron temperature is possible also through the observations of the continuous radiation of the nebula in the visible region of the spectrum (Ref. 44). However, because of the great difficulties connected with the observations of continuous radiation of the nebulae and the comparatively low degree of accuracy of the obtained result, this method for determining T_e was not applied in practice.

3. The Determination of T_e with Respect to the Total Intensity of the Prohibited Lines to the Intensity of the Line H_β (The Method of Energetic Balance)

Initially, the method of energetic balance was proposed by Sobolev (Ref. 10) for the purpose of determining the electron temperature of planetary nebulae. Spitzer (Refs. 11-14) developed a similar method for determining T_e of inter-stellar gas. Baker, Menzel and Aller (Ref. 3) and later Aller (Ref. 7) have developed in a somewhat different form an identical method independent of Sobolev's effort. The principle of the method is the same in all cases; considered are all possible processes during which electrons of the nebula acquire kinetic energy (photo-ionization of hydrogen and helium) and the processes during which this energy is lost (free-free transitions, recombinations, inelastic collisions of ions, etc.).

We have already disucssed this method in detail in the first part of this article. The complete equation for energy balance has the following form /209

$$\bar{\epsilon}_0 = \bar{\epsilon}_{rec} + \epsilon_{ff} + \sum_i \epsilon_i + \epsilon_H , \qquad (40)$$

where $\bar{\epsilon}_0$ is the mean energy of the electrons which are torn out during the ionization of hydrogen and helium (see Table 4) and the $\bar{\epsilon}_{rec}$, ϵ_{ff} and ϵ_H are given respectively by formulae (20), (19) and (18). By following Aller (Ref. 7) it is possible to present the energy, which is lost by the electron on the excitation of the prohibited lines, in the following form

$$\sum_i \epsilon_i = G\ (T_e) \sum_i \frac{I_i}{I_{H_\beta}}\ , \tag{41}$$

where

$$G\ (T_e) = 22.4 \cdot 10^{-20} \frac{b_4\ (T_e)}{\sum_1^{\infty} C_n\ (T_e)} \frac{e^{\frac{9814}{T_e}}}{T_e^{3/2}}\ ,$$

and $\sum_i \frac{I_i}{I_{H_\beta}}$ is the observed total intensity of all the prohibited lines. This intensity is expressed in units of intensity of the H_β lines.

The left term of equation (40) is dependent only on the spectral class of the exciting star and the degree of helium ionization in the nebula. The right term depends on the electron temperature of the nebula, the total intensity of the prohibited lines and the degree of hydrogen ionization in the nebula. The latter is normally on the order of 10^3 to 10^4 and the last term in equation (40) does not have any significant value when all $T_e < 15{,}000°C$. Equation (40) (without the last term) is presented in the form of a nomogram in Fig. 2. The values T_e and $\sum_i I_i/I_{H_\beta}$ are plotted along two extreme scales. The scale in the center contains the value of the parameter $\bar{\epsilon}_0$. The points of intersection of these scales with any straight line yield the values of T_e, $\sum_i I_i/I_{H_\beta}$ and $\bar{\epsilon}_0$ which simultaneously satisfy the equation of balance. Until now the equation for energy balance has been used exclusively as a

method for the determining of the electron temperature. Meanwhile it is obvious from the nomogram that the temperature thus found is extremely sensitive to errors in $\sum_i I_i/I_{H_\beta}$ and $\bar{\epsilon}_0$. The equation for the energy /210 balance is of extremely great interest from the viewpoint of the possibility of determining from observations the values $\bar{\epsilon}_0$ or the mean energy of the ionizing quantum ($h\bar{\nu} = \bar{\epsilon}_0 + xH$) which are dependent on the distribution of energy in the spectrum of a star beyond the Lyman boundary. Actually what do we know about the distribution of radiation beyond the Lyman boundary for such objects, as the nuclei of planetary nebulae? Apparently this type of information might be obtained only either purely theoretically (models), or by some indirect method. The presence of an enormous quantity of neutral hydrogen in the inter-stellar space deprives us forever of the possibility of direct observations of this part of the spectrum of distant objects. Therefore, the possibility of determining the parameter $\bar{\epsilon}_0$ from the equation for the energy balance is of interest not only as a means for checking the values of $\bar{\epsilon}_0$ obtained by theoretical models of stars (Ref. 45), but also from the viewpoint of checking the validity of our concepts concerning the mechanisms of heating and cooling off of the nebula.

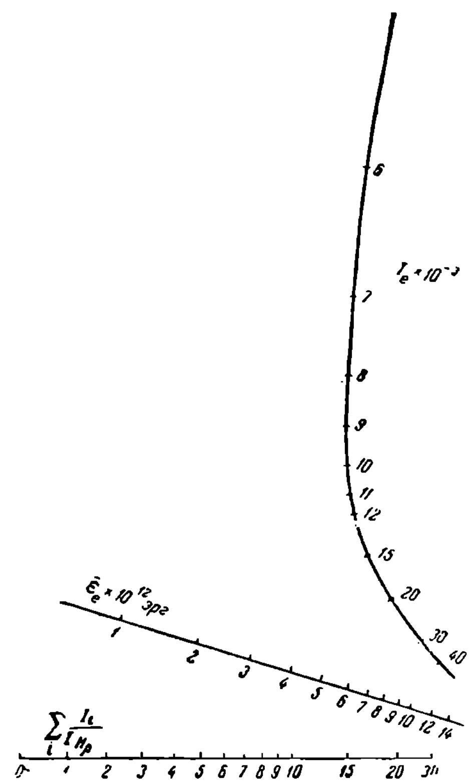


Fig. 2. Nomogram of Energy Balance Equation.

4. The Determination of Electron Temperature of Nebulae by Means of Radio Observations

The visible density of a radio emission flux of the H II region (with a homogeneous electron temperature T_e) located between the observer and the background (with a homogeneous brightness temperature T_b) is expressed as follows:

$$F_{observ} = \frac{2kf^2}{c^2} (T_e - T_b) \int (1 - e^{-\tau})\, d\Omega , \qquad (42)$$

where integration is conducted along the solid angle Ω which occupies the nebula. Here f is the frequency in cycles and τ is the optical thickness in the given frequency. It is obvious from (42) that the electron temperature in the nebula can be easily obtained by way of finding the frequency for which $F_{observ} = 0$ and by measuring T_b within this frequency. Then it is apparent that $T_e = T_b$. As Mills, Little and Sheridan (Ref. 46) have proven the generally accepted value of $T_e \approx$ 10,000°C in this case is almost obvious. However, the method is too sensitive for flux measurements and since the latter, which are calibrated in absolute units, can have an error of ± 20%, and the frequency at which $F_{observ} = 0$, due to the specific noises, is poorly determined, then the method yields in effect not much more than an order of unity for T_e. /211

Thanks to the work by Wade (Ref. 47) a somewhat modified radio method for determining the electron temperature of the H II--regions by the observed densities of the fluxes on two remote radio frequencies has been developed.

After integration by Ω the equation (42) for a spherical nebula acquires the following form:

$$F_{observ} = \frac{4\pi k\theta_0^2 f^2 \,(T_e - T_b)}{c^2} Y(\tau_0), \qquad (43)$$

where

$$Y(\tau_0) = \frac{1}{2} + \tau_0^{-2} [e^{-\tau_0} (\tau_0 + 1) - 1]. \qquad (44)$$

Here θ_0 is the angular dimension of the nebula in radians and τ_0 is the optical depth at the center of the nebula. The value of τ_0 for each frequency can be expressed as a function of the electron temperature. Then the relation of optical depths on the two frequencies f_1 and f_2 can be written as follows

$$\frac{\tau_{0.1}}{\tau_{0.2}} = \frac{\zeta 1}{\zeta 2} \frac{f_2^2}{f_1^2} , \tag{45}$$

where

$$\zeta_{1.2} = 9.70 \cdot 10^{-3} \ln \frac{3kT_e}{2hf_{1.2}} .$$

The further process of finding T_e consists of the following. We substitute the values F and T_b, which are observed on two frequencies; in equation (43), we consider several values of T_e and calculate for these temperatures the functions Y $(\tau_{0.1})$ and Y $(\tau_{0.2})$. By means of formula (44) which can be presented graphically as a curve Y = Y (τ_0), we find the values $\tau_{0.1}$ and $\tau_{0.2}$ and determine for each given T_e the relation $\tau_{0.1}/\tau_{0.2}$. We enter on the graph the relations found for $\tau_{0.1}/\tau_{0.2}$, as a function of the electron temperature. In the same graph is introduced the dependence of $\tau_{0.1}$ and $\tau_{0.2}$ on T_e, which was calculated /212 according to formula (45). The intersection of the two curves gives us the sought value of T_e.

This method is applicable only when τ_0 is not too high. If, on the other hand, the nebula at a given wavelength is not transparent, then it radiates as a black body and its temperature can be determined by measuring the flux of radiation and its angular dimensions. The information in regard to the angular and linear dimensions and also the visible forms of a majority of nearby H II - regions, can be obtained photographically.

5. Determination of the Electron Temperature by Oxygen Lines λ 3727 [O II] and $N_1 + N_2$ [O III]

Equation (28), the principal method for determining electron temperature of planetary nebulae, cannot be in principle applicable to many diffuse nebulae because of the weak ionization of oxygen in them (ions in O III are few or are completely absent). The problem concerning the expediency of using the equation of energy balance as the method of determination of electron temperature has been considered by us in Section 3. As far as the radio method is concerned, although it is applicable to the extensive H II-regions it has very grave short-comings and is entirely unsuitable for many tasks. The principal short-coming of this method is seen directly from formula (45): the observed relation $\tau_{0.1}/\tau_{0.2}$ is very weakly dependent on T_e (as ln T_e). Therefore, the small errors in τ_0 firmly influence the unknown value T_e. Second short-coming is included in the small angular resolution of the radio telescopes.

The absence of a simple, reliable method and, at the same time, one which does not present specific requirements toward observations for the determination of T_e of diffuse nebulae has compelled us to analyze in greater detail the possibilities for determining the temperatures of such nebulae along the bright lines. As a result a method for determining T_e which utilizes the intensity of lines λ 3727 [O II] and $N_1 + N_2$ (Ref. 48).

The comparison of the ionization potentials of the atoms H I and O I (13.595 and 13.614 electron volts, respectively) shows that the degree of ionization of hydrogen and oxygen in regions H II should be identical, i.e., in those regions where hydrogen is luminescent and atoms of neutral oxygen are almost absent. This is also indicated by the absence of a noticeable amplification of the line of the night sky λ 6300 [O II] in the directions toward diffuse nebulae.

On the other hand, a comparison of the ionization potentials He II and O II (54.403 and 54.93 electron volts, respectively) and the absence in diffuse nebulae of ions of He III (the line λ 4686 He II is not observed) indicates that the ions of O IV and higher stages of oxygen ionization are also absent. Thus, oxygen in diffuse nebulae (and also in planetary nebulae with low excitation) is found only in two ionization levels - O II and O III. This gives us one more method of determining the electron temperature. This concept is included in the following. If it is considered that the relative chemical composition of all nebulae is identical, then the following should be valid

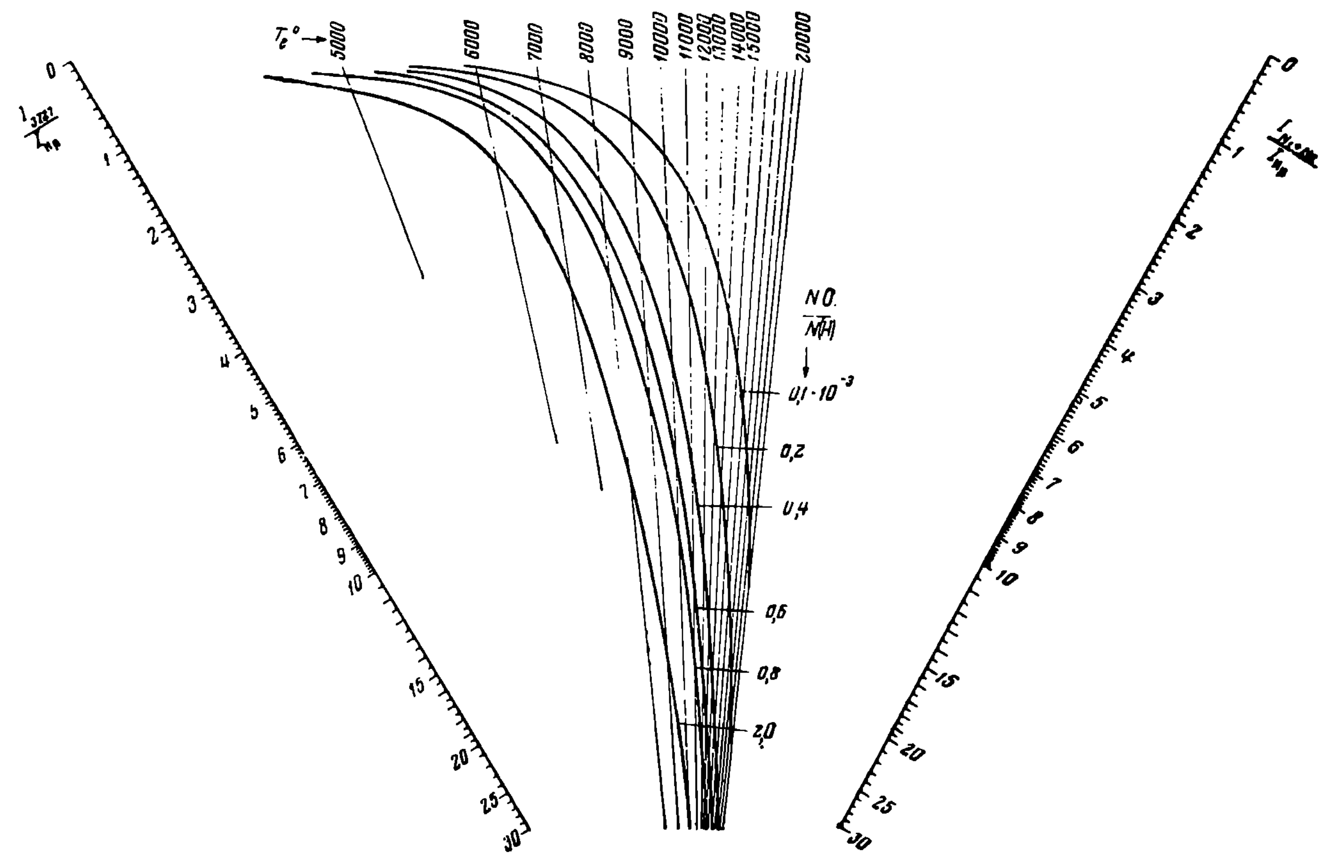


Fig. 3. Nomogram of Equation (47). The electron temperature of the nebula is determined when N(O)/N(H) and the observed values I_{3727}/I_{H_β} and $I_{N_1+N_2}/I_{H_\beta}$.

$$\frac{N\ (O\ II)}{N\ (H)} + \frac{N\ (O\ III)}{N\ (H)} = \frac{N\ (O)}{H\ (H)}\ , \tag{46}$$

where N (O)/N (H) is the total number of oxygen atoms per unit of volume of the nebula relative to the number of hydrogen atoms. Now if the relations N (O II)/N (H) and N (O III)/N (H) are expressed through the intensities of the lines λ 3727, $N_1 + N_2$ and H_β, then we will obtain the equation

$$\frac{I_{3727}}{I_{H_\beta}} \cdot \frac{1}{\Theta_{[O\ II]}} + \frac{I_{N_1+N_2}}{I_{H_\beta}} \frac{1}{\Theta_{[O\ III]}} = \frac{N\ (O)}{N\ (H)}\ , \tag{47}$$

where Θ for a given line is the function of only the temperature:

$$\Theta = 0.38 \cdot 10^{14} \frac{\Omega h\nu}{\varpi_A} \frac{T_e \cdot \theta\ (T_e)}{b_4\ (T_e)}\ e^{-\frac{9414}{T_e} - \frac{x_{AB}}{kT_e}}\ . \tag{48}$$

Thus, the entire process of determining the temperature is reduced to a selection of a value of T_e where the functions of together with the observed intensities of the lines λ 3727 and $N_1 + N_2$ at a given magnitude of N (O)/N (H) (identical for all nebulae) would satisfy equation (47). For a quick calculation of such a temperature a nomogram of this equation is constructed (Fig. 3). By combining the points of two extreme scales on which are placed the observed intensities of the lines λ 3727 and $N_1 + N_2$ by a straight line, it is possible to find the value T_e which satisfies equation (47) for any given N (O)/N (H). As is obvious from the nomogram, the electron temperature which was found in this way will be significantly dependent on the values of N (O)/N (H) which we have obtained. If the relative content of oxygen were accurately known, equation (47) would yield the exact value of T_e. However, various methods and different authors give diverse values for N (O)/N (H). It is, therefore, necessary to take the average for all determinations (see Table 5). This method makes it possible to have a single temperature scale for all nebulae although its zero-point can be somewhat shifted to one side or another relative to the true temperature scale.

REFERENCES /215

1. W. H. Wright. Publications of the Lick Observatory, 13, 193, 1918.
2. V. A. Ambartsumyan. Teoreticheskaya Astrofizika (Theoretical Astrophysics) Gosudarstvyennoye Tekhnicheskoye Izdatel'stvo, 1939.
3. J. Baker, D. Menzel and L. Aller. Astrophysical Journal, 88, 422, 1938.
4. M. Hebb and D. Menzel. Astrophys. J., 92, 408, 1940.
5. D. Menzel, L. Aller and M. Hebb. Astrophys. J., 93, 230, 1941.
6. D. Menzel and L. Aller. Astrophys. 94, 30, 1941.
7. L. Aller. Astrophys. J., 118, 547, 1953.
8. J. Baker and D. Menzel. Astrophys. J. 88, 52, 1938.
9. A. Burgess. Monthly Notices of the Royal Astronomical Society 118, 477, 1958.
10. V. V. Sobolev. Uchenyye Zapiski Leningradskogo Universiteta, No. 82, 1941.
11. L. Spitzer. Astrophys. J. 107, 6, 1948.
12. L. Spitzer. Astrophys. J. 109, 337, 1949.
13. L. Spitzer and M. Sayedov. Astrophys. J. 111, 593, 1950.
14. L. Spitzer. Astrophys. J. 120, 1, 1954.
15. H. C. van de Hulst. Rech. Utrecht, 11, 1, 1946.
16. A. Underhill. Publications of the Dominion Astrophysical Observatory, Victoria, B. C. 8, No. 12, 1951.
17. A. Underhill. Kobenhavns Meddelanden, No. 151, 1950.
18. A. Underhill. Etoiles raies d'emission. University de Liege. Belgique, 1958, p. 17.
19. J. C. Pecker. Annales d'Astrophysique, 13, 433, 1950.
20. G. Traving. Zeitschrift für Astrophysik, 36, 1, 1955.

REFERENCES (Cont'd)

21. G. Traving. Z. f. Astrophys., 41, 215, 1957.

22. S. Saito and A. Uesugi. Publ. Astr. Soc. Japan, 11, 90, 1959.

23. B. Stremgren. Astrophys. J. 89, 526, 1939.

24. R. Ye. Gershberg. Izvestiya Krymskoy Astrofizicheskoy Observatorii, 26, 324, 1961.

25. L. H. Aller and R. Minkowski. Astrophys. J. 124, 110, 1956.

26. L. H. Aller and W. Liller. Astrophys. J. 130, 93, 1960.

27. L. H. Aller. Astrophys. J. 125, 84, 1957.

28. L. H. Aller and D. H. Menzel. Astrophys. J. 102, 239, 1945.

29. K. Hunger. Z. f. Astrophys., 49, 129, 1960.

30. S. B. Pikelner. Izv. Krymsk. Astrofiz. Observ. 12, 93, 1954.

31. J.W. Chamberlain. Astrophysical J. 117, 387, 1953.

32. M. J. Seaton and D. E. Osterbrock. Astrophys. J. 125, 66, 1957.

33 M. Seaton. Kosmicheskaya Gazodinamika (Cosmic Gas Dynamics), Izd-vo Inostrannoy Literatury, 1960, p. 121.

34. R. H. Garstang. Monthly Notices Roy. Astron. Soc., 111, 115, 1951.

35. M. J. Seaton. Monthly Notices Roy. Astron. Soc., 114, 154, 1954.

36. K. H. Böhm. Astrophys. J., 123, 379, 1956.

37. T. Page. Monthly Notices Roy. Astron. Soc., 96, 604, 1936.

38. T. Page. Astrophys. J., 96, 78, 1942.

39. D. Barbier. Annales d'Astrophysique, 7, 80, 1944.

40. A. Ya. Kipper. O Razvitii Sovietskoy Nauki v Estonskoy SSR (On the Development of Soviet Science in the Estonian SSR) 1940-1950. Estonskoye Gosudarstvennoye Izdatel'stvo 1950, p. 316.

41. L. Spitzer and J. L. Greenstein. Astrophys. J., 114, 407, 1951.

REFERENCES (Cont'd)

42. H. Andrillat. Ann. d'Astrophys., Suppl., No. 1, 1955.

43. M. J. Seaton. Ann. d'Astrophys., 17, 296, 1954.

44. M. J. Seaton. Monthly Notices Roy. Astron. Soc., 115, 279, 1955.

45. V. I. Pronik. Astron. Zhur. 37, 1001, 1960.

46. B. Y. Mills, Little, J. Sheridan. Australian Journal of Physics, 9, 218, 1956.

47. C. M. Wade. Austr. J. Phys., 2, No. 3, 1958.

48. V. I. Pronik. Izv. Krymsk. Astrofiz. Observ., 17, 14, 1957.

CHRONICLES

CONFERENCE ON EXTRAGALACTIC ASTRONOMY AND COSMOLOGY

A conference of the Commission on Cosmogony from the Astronomic Council of the Academy of Sciences of the USSR dedicated to extragalactic astronomy and cosmology took place on June 25-28, 1961 in Moscow at the P. K. Schternberg State Astronomic Institute. The conference was organized for the purpose of preparation for the symposium on the subjects adopted for the 11th Congress of the International Astronomic Union.

The meeting on June 27th, dedicated to the theoretical problems of cosmology, was conducted in the Physical Faculty Department of Moscow State University jointly with participants of the All-Union Conference on Gravitation.

In V. A. Ambartsumyan's lengthy report (Byurakan Astronomical Observatory) basic problems of extragalactic investigations were considered (the report is published in its entirety in the present volume of the collection "Problems of Cosmogony").

B. A. Vorontsov-Velyaminov (Schternberg State Astronomic Institute) focused his main attention on the inter-acting galaxies and problems of the physical interpretation of their observed interaction (report is being published).

Yu. P. Pskovskiy (Schternberg State Astronomic Institute) dedicated his address to the investigation of distances, motions and distribution of galaxies in a sphere with a radius of 15 megaparsecs (report is being published).

E. A. Dibay (Schternberg State Astronomic Institute) proposed the interpretation of an asymmetric distribution of gas and dust in spiral galaxies as a result of the joint influence of rotation and radial motion along the limb.

N. S. Kardashev (Schternberg State Astronomic Institute) reported on the study of the deviation of the surface maximum density of hydrogen from the plane of symmetry of the galaxy and the interaction of the gas of the galaxy with intergalactic gas (report is being published).

T. A. Agekyan and I. V. Petrovskaya (Leningrad State University) discuss the questions concerning the stationarity of a cluster of galaxies in Coma Berenices and the distribution of density in the cluster connected with this.

A. T. Kallogyan (Byuratan Astronomical Observatory) discussed the dynamic instability of certain groups of galaxies and of a cluster in the Corona Borealis.

In the report by Ya. A. Smorodinskiy and B. M. Pontekorovo /217
Institute of Nuclear Research (OI Ya. I, Dubna) the potential role of neutreno in astrophysics and cosmogony was considered.

V. L. Ginzburg (Institute of Physics, the Academy of Sciences of the USSR) gave an account of the investigations of cosmic rays which are shedding light on the nature of radio galaxies and the Metagalaxy.

L. A. Frank-Kamenetskiy (The Institute of Anthropology and Ethnography of the Academy of Sciences of the USSR) discussed the hypotheses concerning the plural birth of nucleon pairs by thermic photons in an exposed cosmological model.

E. M. Lifshitz, I. M. Khalatnikov and V. V. Sudakov (The Institute of Physical Problems of the Academy of Sciences of the USSR) on the basis of research on the characteristics of the cosmological solutions to the equations of gravitation have shown that the existence of the physical property of time is not a characteristic of the cosmological models in the general theory of relativity.

A. L. Zelmanov (Schternberg State Astronomic Institute) reported on the theory which he is developing on an antisotropic inhomogeneous universe.

M. F. Shirokov (Moscow Aviation Institute) and I. Z. Fisher dedicated their address to the theory of an inhomogeneous isotropic universe.

A. Ya. Kipper (Institute of Physics and Astronomy of the Academy of Science of the Estonian SSR) discussed the gravitational paradox in Newton's theory (report is being published).

E. L. Zelmanov (Schternberg State Astronomic Institute) considered the gravitational paradox and the quasi-Newton approximation.

I. D. Novikov (Schternberg State Astronomic Institute) gave an account concerning some cosmological mosels in a quasi-Newton approximation.

G. M. Idlis (The Astrophysical Institute of the Academy of Sciences of the Kazakh SSR) reported on research conducted by him in conjunction with R. Kh. Gainullina and Z. Kh. Kurmakayev on the search for visible compressions of distant spherical components of multiple galaxies due to the Einstein effect.

L. M. Ozernoy (Shternberg State Astronomic Institute) considered some problems of gravitational condensation of galaxies and globular clusters.

G. I. Naan (The Academy of Sciences of the Estonian SSR) dedicated his address to a critique of certain philosophical works in which the "cosmological infinity" concept is acceptably simplified or distorted and showed how it is necessary to consider this concept from the point of view of contemporary physics.

About 100 specialists, physicists and astronomers participated in the conference. Nearly all the reports were subjected to lively discussion.

Ye. L. Ruskol

Translated by Joseph L. Zygielbaum
Electro-Optical Systems, Inc.
Pasadena, California

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Technical information generated in connection with a NASA contract or grant and released under NASA auspices.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

TECHNICAL REPRINTS: Information derived from NASA activities and initially published in the form of journal articles.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities but not necessarily reporting the results of individual NASA-programmed scientific efforts. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION DIVISION

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546